HISTONE MODIFICATION, GENE REGULATION AND EPIGENETIC MEMORY IN EMBRYONIC STEM CELLS

By

Elsa Boudadi

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Chromatin and Gene Expression Group,
Institute of Biomedical Research,
College of Medical and Dental Sciences,
The University of Birmingham
July 2011

Abstract

Histone modifications are thought to act as a layer of epigenetic information, because of their strong association with gene expression, and their potential role in transcriptional memory. However, although specific histone modifications correlate with transcriptional status, whether they play a causative role or act in the long term inheritance of gene expression patterns is unclear.

In order to explore this, the histone deacetylase inhibitor valproic acid (VPA) was used to induce hyperacetylation of histones in embryonic stem cells. Surprisingly, although global levels of acetyl marks were highly increased by VPA treatment (up to 16-fold), only 10% of genes showed transcriptional changes. Interestingly, these global changes in histone modification were not reflected in the changes at individual genes where increases in acetylation were rarely greater than 2-fold. Furthermore, changes in acetylation levels did not correlate with transcriptional effects. Wash-out experiments showed that transient VPA treatment could not induce long term effects on transcription, even during ES cell differentiation when histone modifications play a crucial role. Finally, the role of polycomb silencing in the response to VPA treatment was assessed using an ES cell line in which the polycomb components Eed and Ring1b had been knocked out. Target genes showed small up-regulation in knockout cells but VPA did not further induce transcription.

It was concluded that histone acetylation plays an important role in transcription but additional signals are required for transcriptional induction and cellular memory. My results suggest the existence of protective mechanisms against hyperacetylation and highlight the complexity of epigenetic regulation, potentially involving many layers of control.

Acknowledgments

I would like to sincerely thank my supervisor Professor Bryan Turner for the invaluable opportunity he offered me. I benefited during my PhD from constant support, trust and freedom in the lab, from a challenging but always friendly environment as well as many opportunities to learn about epigenetics at various conferences.

I also would like to thank Dr. Karl Nightingale for the tremendous help with the writing of this manuscript, for being always available for me and for teaching me perfect and only useful English. Many thanks also for welcoming me and helping me to find a place to live when I arrived four years ago.

I'm very grateful to Dr. Laura O'Neil for the immense help with ES cell culture and the guidance in the lab and in the designing of experiments. Many thanks as well for the spontaneous help with breakfast seminars and other stressful situations.

Thanks to all the past and present members of the lab for being such good colleagues and friends. Special thanks to Lesley, Edith, Hannah, Marianne, Rebecca, Richard and John for his useful advice during the writing of this manuscript.

Thanks to all the wonderful people I met here and that made my life in England so exceptional. I will deeply miss Nil and Marcin, Anna, Philipp, Annelise and Romuald, Edith and Carl, Paulo, Sara, Ian, Katia, Jan, Lizzy, Clare and Sarah.

Special thanks to Allan who was the best housemate one could ever wish to have. Thanks also for starting your PhD six months before me and warning me of all the good and less good things to come.

Finally, I would like to thank my friends and family for the unconditional support and love and for making every holiday in Paris so magical. Immense thanks to my grandmother for keeping in touch with the modern world and allowing long and warm conversation on skype.

Table of content

Table of Figures

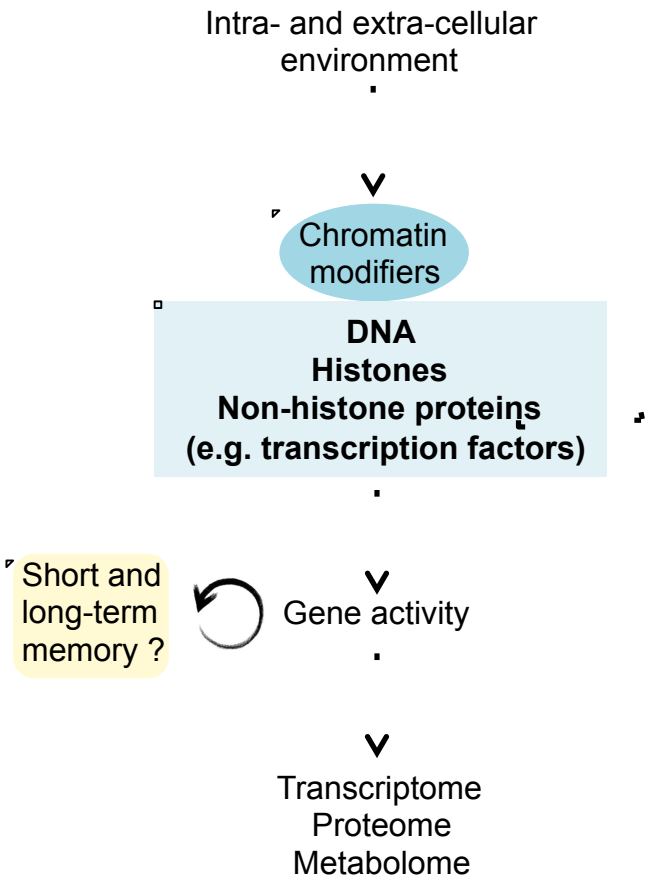
Table of tables

1. Introduction

1.1 Chromatin sensing of the environment

The development of a complex organism requires a single genome to generate a multitude of cell phenotypes. The term "*epigenetic"* appeared in 1942 along with the idea that genetics and developmental biology should be related. Conrad Waddington, by using the prefix epi- ("above" in Greek), suggested the existence of mechanisms besides the gene sequence that participate to the acquisition and maintenance of a specific cell identity (Holliday, 2006). In addition to differentiation processes, a cell will be exposed to multiple events including mitotic and apoptotic signals, DNA repair, metabolite availability, drug exposure or the general chemical environment. All those events require the more or less rapid activation or silencing of specific genes, and for these changes to be maintained for the short or long term (i.e. more than one cell cycle). Thus, the cell has to sense its environment in order to adapt transcription levels of important genes and maintain proper function (Bird and Tweedie, 1995). In the nucleus, the DNA is associated with histone proteins, RNAs and non-histone proteins, in an organised structure named "the chromatin fibre". Chromatin components are subject to chemical modifications, which are then interpreted by the transcriptional machinery for appropriate gene expression. Importantly, the activities of enzymes that add and remove those chemical "marks" have been shown to directly or indirectly depend on the environment. It has been proposed that the chromatin fibre, via modifying enzymes, acts as an environmental sensor and a signalling unit for the alteration of gene expression **(Figure 1.1)**. Thus, it is of general interest to understand how environmental agents can impact on the activity of the genome in a short or long-term manner (Turner, 2009).

Figure 1.1 Role for the chromatin fibre as an environmental sensor in short and long-term establishment of gene expression



Gene expression activity is regulated by a chromatin network composed of histones, DNA and non-histone proteins including transcription factors. Those components influence transcription through their modification by groups of highly specialised enzymes. The activity of modifying enzymes can be altered by environmental factors like inhibiting/activating molecules or metabolite availability. Thereby, gene expression is potentially affected by external signals. If chromatin modifications are causative of transcription and are inherited through cell cycle, an environmentally-induced epigenetic change could impact on transcription for several rounds of cell division.

1.2 Chromatin fibre and modifications

1.2.1 Chromatin organisation and dynamics

The defining characteristic of eukaryotic cells is the presence of a nucleus. It is in this organelle that the bulk of genetic information is maintained. The DNA molecule can reach two meters in length and the nucleus is only a few microns in diameter. For this reason, the DNA fibre undergoes different levels of compaction (from the naked DNA to the metaphase chromosome) with the help of associated proteins. The complex resulting from the association between the DNA and the highly conserved histone proteins was named the chromatin fibre and was described as an array of nucleosomes. The first function of a nucleosome is therefore to package the DNA in order to resolve steric effects but also to regulate DNA accessibility, an important factor for gene expression. A nucleosome consists of a nucleosome core, the linker DNA, and associates with the linker histone H1 (Luger et al., 1997).

The nucleosome core is formed by an octamer of histone proteins around which 146 base pairs of DNA wrap in a left-handed superhelix. The best representation of the nucleosome to date was obtained using high-resolution X-ray crystallography **(Figure 1.2)**. The core histone octamer contains two copies of each histone H2A, H2B, H3 and H4. Two H3-H4 pairs form a tetramer through the interaction between the two H3 molecules. Two H2A-H2B dimers then interact with the H3-H4 tetramer through a 4-helix bundle between H2B and H4 (Luger et al., 1997).

Figure 1.2 Crystal structure of the nucleosome core

High resolution X-ray crystallography allowed visualisation of the nucleosome structure in 1997. The histone core octamer results in the association of a 2x(H3-H4) tetramer and two (H2A-H2B) dimers. 146 base pairs of DNA wrapped around the histones to form the nucleosome. Importantly, N-terminal tails of histones that extend out of the nucleosome core are not visible due to their unstructured nature; Two strands of DNA (green and orange), H4 (green), H3 (blue), H2A (yellow), H2B (red); Face on view (left), side on view (right).
Taken from Luger et al., 1997.

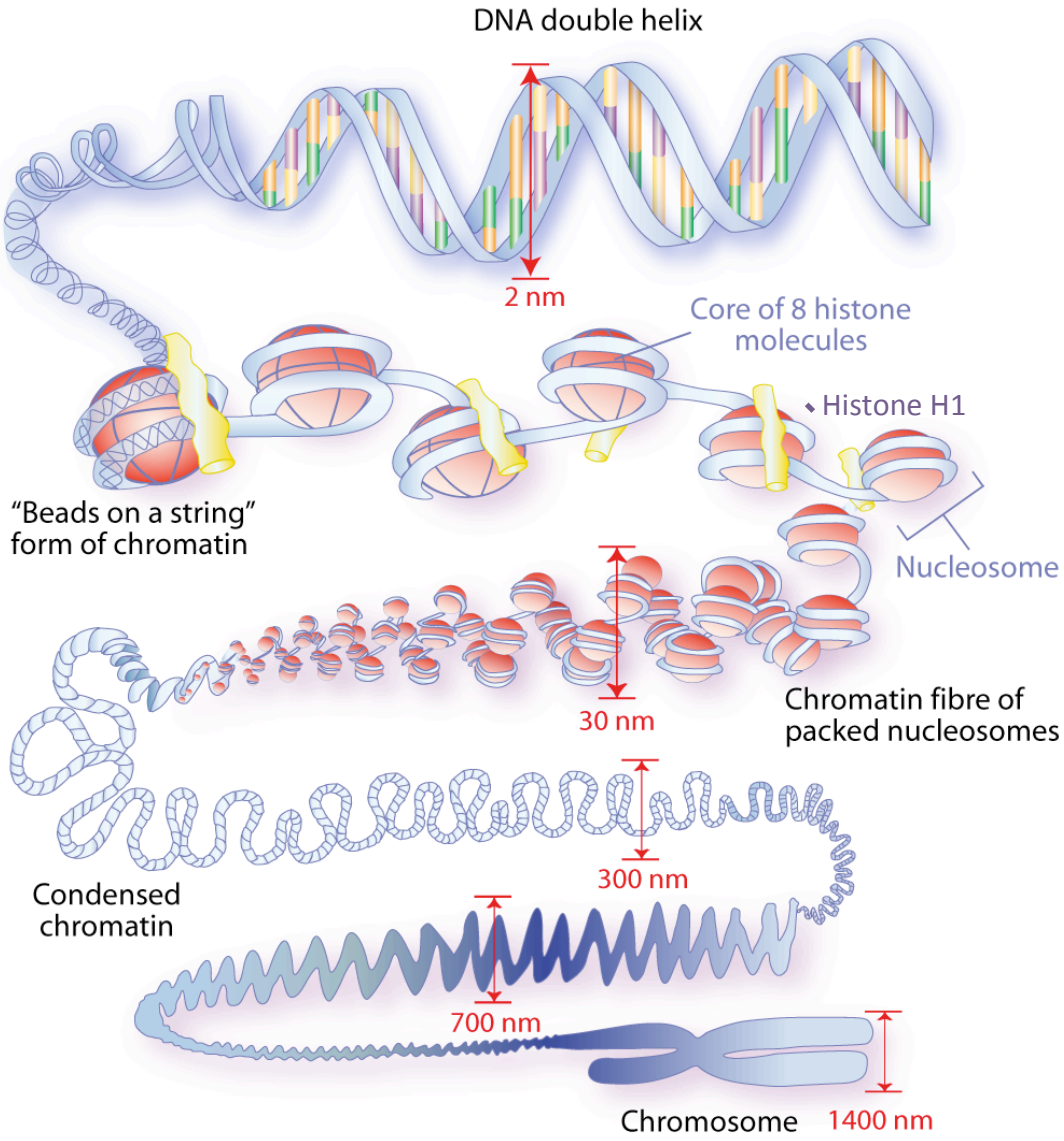
A fifth histone protein, the lysine-rich histone H1, was found to sit on top of the nucleosome outside the core octamer and to associate with the DNA at the entry/exit sites of the nucleosome. H1 was found to play a role in the stabilisation of compact chromatin fibre as an H1-depleted fibre could still fold into some higher order structures in the presence of NaCl (Hansen et al., 1989). However, H1 was found to be indispensible for proper mouse development (Fan et al., 2003) and for the compaction of metaphase chromosomes (Maresca and Heald, 2006).

Finally, the portion of free DNA between adjacent nucleosomes was called the linker DNA. The length of linker DNA partly dictates nucleosome positioning and is considered as an important factor in chromatin function. In fact, it is likely that the position of nucleosomes in the chromatin fibre is a key determinant for the accessibility of various DNA-binding proteins and therefore affects gene regulation. For instance, a high resolution study in yeast showed that gene expression correlated inversely with nucleosome occupancy around the transcription start site (Lee et al., 2007c). The insulator-binding protein CTCF was also found to position nucleosomes around its binding site (Fu et al., 2008). What determines nucleosome positioning is still under debate and the DNA sequence, chromatin remodelling enzymes, histone modifications and transcription factors binding have been implicated (Arya et al., 2010).

The chromatin fibre is a dynamic structure that is associated with different levels of compaction **(Figure 1.3)**. The negatively charged “naked DNA” associates with the positively charged histone proteins into nucleosomes linked to each other by a portion of linker DNA. At low salt concentrations, electron microscopy revealed this array as a “beads on a string” structure (Olins and Olins, 1974). However, when the ionic concentration was raised, reducing DNA-DNA repulsion, the chromatin was shown to fold into more compacted secondary structures, including the “30 nm fibre” (Thoma et al., 1979) and the 300nm fibre (Widom, 1986). Finally, a further increase in Mg^{2+} concentration was shown to induce the 300nm fibres to interact with each other side-by-side and end-to-end (Widom, 1986). How higher order chromatin structures are formed and stabilized remains unclear. Histone H1 phosphorylation, topoisomerase II and condensin complexes have been implicated in metaphase chromosome formation but mechanisms have not been fully described (Belmont, 2006).

Early light and electron microscopy studies identified two types of chromatin in the nucleus. The heterochromatin was shown to remain compact throughout the cell cycle and the euchromatin was described as less condensed during interphase. The heterochromatin was therefore traditionally defined as the silenced regions of the genome in opposition to the open-to-transcription euchromatin. The picture became more complex when the heterochromatin was divided into two types, constitutive and facultative heterochromatin, where constitutive heterochromatin usually corresponds to specific regions such as centromeres or telomeres. On the other hand, regions of facultative heterochromatin vary between cell types and are regulated, for instance, by differentiation processes (Dimitri et al., 2005). The inactive X-chromosome in

Figure 1.3 The different levels of chromatin compaction



The chromatin fibre is not a static structure and can be more or less compacted, as one of its major functions is to package the DNA into the nucleus and regulate binding protein access. Although chromatin is very dynamic, several layers of chromatin structure have been identified including the "beads on a string" fibre, the 30nm, 300nm and 700nm fibres as well as the very compacted metaphase chromosome. Histone H1 binds to the external face of the nucleosome in order to stabilise the 30nm fibre. However, how high levels of compaction are established is not fully understood yet.

female mammals would be one example of facultative heterochromatin. It is important to note that transcription is not a constant process so specific regions of chromatin are unlikely to stay in a unique conformation but must constantly adapt to cellular needs (Dimitri et al., 2009).

1.2.2 Chromatin modifications

A series of modifications to the DNA itself and the histone proteins have been characterised and are termed “epigenetic modifications”. Many studies have correlated these “marks” with gene activity, and several mechanisms describe how specific modifications induce transcriptional activation or silencing. Thus, it seems that the chromatin is not a simple packaging tool but also a platform for signalling that regulates gene expression.

1.2.3.1 Histone modifications

All histone proteins possess unstructured N-terminal tails that extend out from the nucleosome core. These exposed tails are subject to a wide variety of post-translational modifications such as acetylation, methylation, phosphorylation, ubiquitination or ADP-ribosylation **(Figure 1.4.a)**. Specific histone modifications are deposited on specific residues. For instance lysines (K) can be acetylated, ubiquitinated or sumoylated, while serines (S) are phosphorylated. Methylation occurs on both lysines and arginines (R) and at different levels as a residue can be mono-, di- or trimethylated (Kouzarides, 2007). Histone modifications alter transcription via direct or indirect effects. For instance, acetylation was shown to loosen the chromatin fibre by neutralising the positive charge of lysines. Thereby the

association of histones with DNA is weakened, allowing transcription factors and RNA polymerase to bind regulatory sequences (Tse et al., 1998). Histone modifications can also indirectly lead to changes in gene expression as they are specifically recognised by proteins that will then impact on the function of the adjacent chromatin (Taverna et al., 2007) **(Figure 1.5.)**.

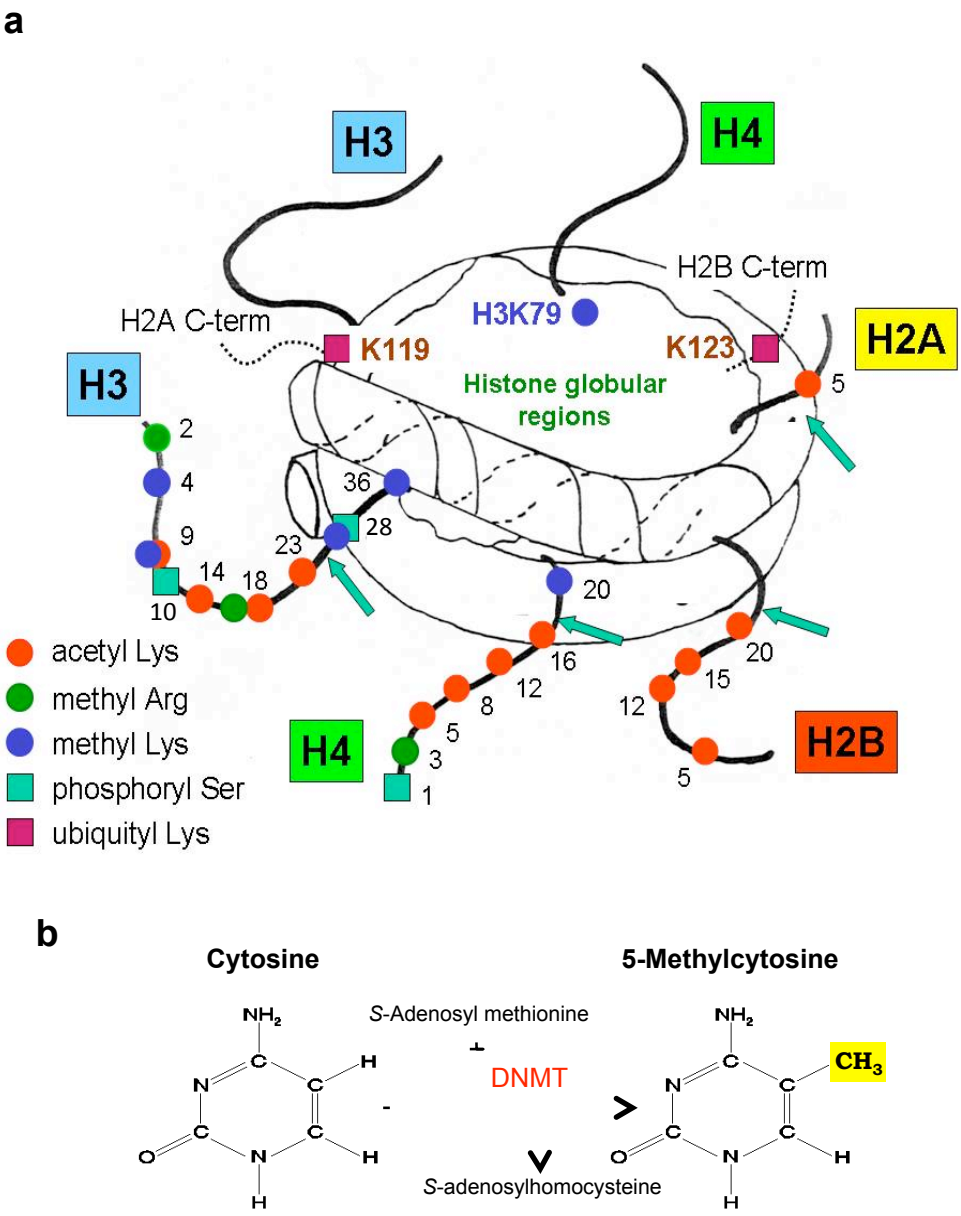
Mass spectrometry identified more than 30 modifications in the core domains of histones. Although their roles is still unknown, they are suspected to impact on chromatin structure and transcription, for instance by altering the interaction between histones and DNA (Mersfelder and Parthun, 2006).

1.2.3.2 DNA methylation

Another major epigenetic modification, DNA methylation, has been consistently associated with transcriptional silencing. In eukaryotes, addition of a methyl group occurs on the cyclic carbon 5 of cytosine, predominantly at symmetrical CpG dinucleotides **(Figure 1.4.b)**. However, evidence of non-CpG methylation was found in embryonic stem cells (Ramsahoye et al., 2000). The distribution of DNA methylation in the genome is uneven, as significant numbers of short regions of DNA are highly enriched in CG and often unmethylated when compared to bulk DNA. These so-called “CpG islands” mark the 5’ end of genes as they generally cover a portion of the promoter and the first exon. A little less than half of all genes are associated with CpG islands in mouse, and are either housekeeping genes or tissue restricted (Cross and Bird, 1995). Although CpG islands remain mostly unmethylated regardless of transcription activity, some important exceptions including methylation of the inactive X-chromosome or CpG islands associated with imprinted genes

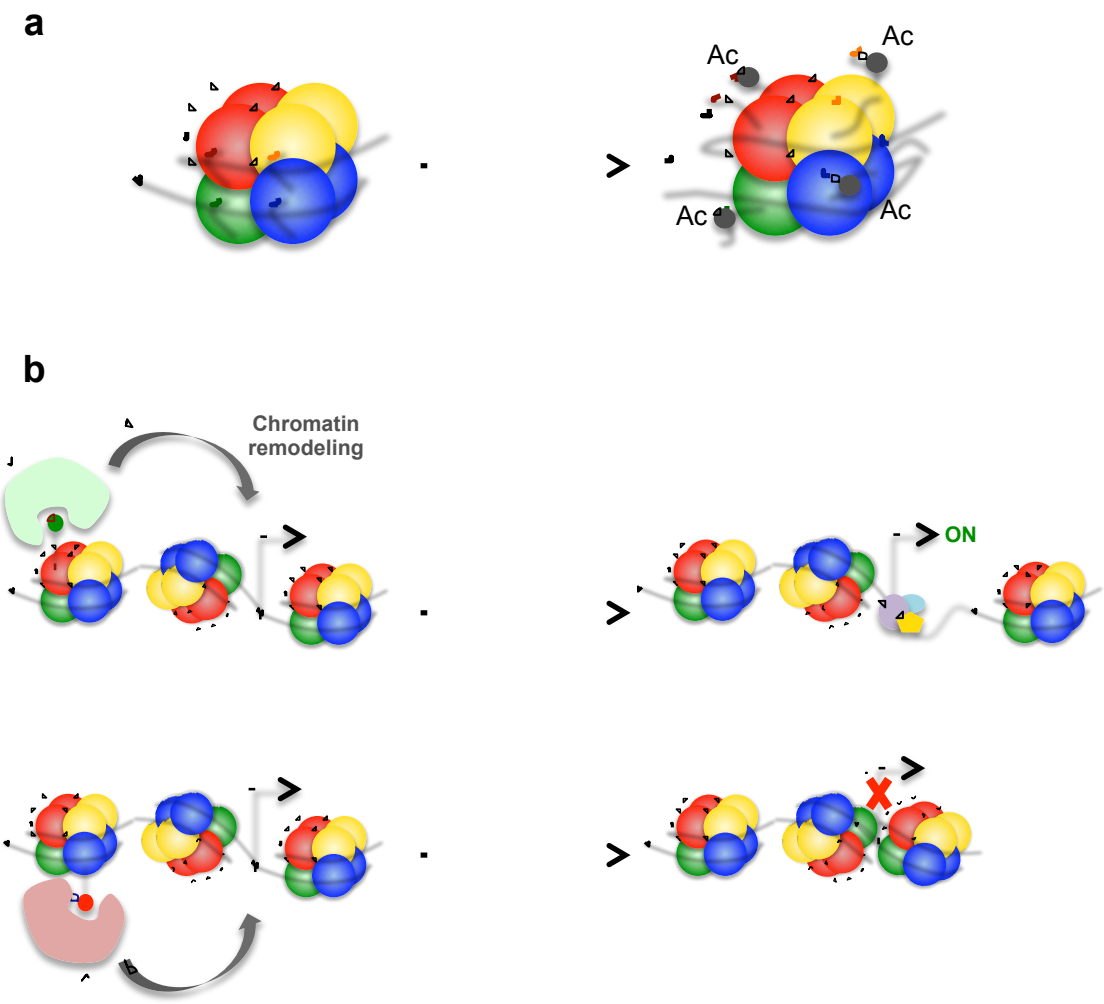
(Ideraabdullah et al., 2008). Like histone modifications, DNA methylation regulates transcription both directly or indirectly and is generally associated with silencing. For instance, methylated CpG is able to block the binding of certain transcription factors like c-Myc or E2F (Campanero et al., 2000). In addition, a family of proteins that contain a methyl-CpG-binding domain (MBD proteins) play a role in silencing by targeting chromatin remodelling complexes to regulatory regions (Lopez-Serra and Esteller, 2008).

Figure 1.4 Modifications of histone tails and deoxy-cytosines



a. Histone tails are the subjects of many post-translational modifications that occur on specific residues. Lysines can be acetylated, methylated or ubiquitinated. Phosphorylation occurs on serines and methylation on both lysines and arginines. Although modifications are shown only on one copy, they can take place on both copies of each histone.
Adapted from Spotswood and Turner, 2002.
b. DNA can be modified and more specifically methylated. The Majority of methylation occurs at cytosine residues and is catalysed by DNA methyltransferases (DNMTs).

Figure 1.5 Direct and indirect effects of histone modification



Histone modifications affect transcription via direct or indirect effects. (**a**) Acetylation was shown to loosen the chromatin fibre by neutralising the positive charge of lysines. As a result, the association of histones with DNA is weaker and transcription factors or the RNA polymerase can bind to regulatory sequences. (**b**) Histone modifications can also indirectly change gene expression when being specifically recognised by proteins that will for instance promote chromatin remodelling.

1.2.3.3 Histone variants

The incorporation of histone variants, as opposed to the canonical histones (H2A, H2B, H3 and H4), into the chromatin fibre is an additional well-established epigenetic phenomenon. Eukaryotic cells contain variants of histone H3 and H2A and an essential characteristic of those variants resides in their ability to distinguish chromatin states (Jin et al., 2005). For example, H2A.X is phosphorylated at DNA double-strand breaks sites while H2A.Z and H3.3 are found in regions of active transcription. Furthermore, MacroH2A was involved in heterochromatin formation and the histone variant CENP-A was shown to localized at centromeres (Jin et al., 2005). Mechanisms of histone variant incorporation at specific sites are now well studied as they could provide an explanation for the inheritance of expression patterns (Henikoff et al., 2004).

1.2.3.4 Non-coding RNA

Finally, being not directly associated with genomic variation, the mechanisms by which non-coding RNA (ncRNA) regulates transcription are also classified as an epigenetic phenomenon. It is now clear that a vast proportion of the mammallian genome does not code for proteins. However, the idea that non protein-coding portions of DNA are important regulatory regions is becoming well accepted. Many studies have provided evidence for a role of ncRNA in mRNA degradation and regulation of translation as well as X-chromosome inactivation or gene imprinting. Mechanisms are not fully understood yet, but specific ncRNA were recently shown to recruit chromatin modifiers like DNA methyl transferases (Zhou et al., 2010).

1.3 Histone modifications and transcription

1.3.1 Deposition of histone modifications

Large families of specific enzymes are responsible for the deposition or removal of histone modifications. The enrichment of each mark is therefore the result of a dynamic process involving the activity of two enzymes that catalyse opposite reactions (**Table 1.1 and 1.2**).

1.3.1.1 Role of enzymes

Acetyl groups are deposited on lysine residues by histone acetyl-transferases (HAT) and removed by histone deacetylases (HDAC). HATs and HDACs have been classified into different groups according to sequence homology or biological function. Two groups of HAT, A (nuclear) and B (cytoplasmic), have been classified into at least three families: The “GNAT” (including the cytoplasmic *Hat1,* the yeast member *Gcn5* and its human homologue *PCAF*), the “MYST” (including *MOZ*, *ybf2/Sas3*, *Sas2* and *Tip60*) and the “p300/CBP” (human paralogues) families. All HATs use Acetyl-CoenzymeA (CoA) as a donor for the transfer of acetyl residues. Interestingly, Gcn5/PCAF contains a bromodomain responsible for the recognition of the acetyl-lysine and some members of the MYST family contain a chromodomain that binds to non-coding RNA. Finally, the different HATs are not highly specific but do show preferential activity for certain histone isoforms and lysine residues (Marmorstein, 2001).

Table 1.1 Acetylation and methylation of lysine residues

New nomenclature	Human	Drosophila	S. cerevisiae	Targeted residue
Histone acetyltransferase				
KAT1	HAT1	CG2051	Hat1	H4K5,12
KAT2		dGCN5/PCAF	Gcn5	H3K9,14,18,23,36; H2B; yHtzl
KAT2A	hGCN5			H3K9,14,18; H2B; H4K8
KAT2B	PCAF			H3K9,14,18; H2B
KAT3		dCBP/NEJ		H4K5,8; H3K14,18
KAT3A	CBP			H2AK5; H2BK12,15
KAT3B	P300			H2AK5; H2BK12,15
KAT4	TAF1	dTAF1	Taf1	H3; H4
KAT5	TIP60/PLIP	dTIP60	Esa1	H4K5,8,12,16; H2A
KAT6		CG1894	Sas3	H3K14,23
KAT6A	MOZ/MYST3	ENOK		H3K14
KAT6B	MORF/MYST4			H3K14
KAT7	HBO1/MYST2	CHM		H4K5,8,12; H3
KAT8	HMOF/MYST1	dMOF	Sas2	H4K16
KAT9	ELP3	dELP3/CG15433	Elp3	H3
KAT10			Hap2	H3K14; H4
KAT11			Rtt109	H3K56
KAT12	TFIIIC90			H3K9,14,18
KAT13A	SRC1			H3; H4
KAT13B	ACTR			H3; H4
KAT13C	P160			H3; H4
KAT13D	CLOCK			H3; H4
Histone methyltransferase				
KMT1		Su(Var)3-9		H3K9
KMT1A	SUV39H1			H3K9
KMT1B	SUV39H2			H3K9
KMT1C	G9a			H3K9
KMT1D	EuHMTase/GLP			H3K9
KMT1E	ESET/SETDB1			H3K9
KMT1F	CLL8			
KMT2			Set1	H3K4
KMT2A	MLL1	Trx		H3K4
KMT2B	MLL2	Trx		H3K4
KMT2C	MLL3	Trr		H3K4
KMT2D	MLL4	Trr		H3K4
KMT2E	MLL5			H3K4
KMT2F	hSET1A			H3K4
KMT2G	hSET1B			H3K4
KMT2H	ASH1	Ash1		H3K4
KMT3			Set2	H3K36
KMT3A	SET2			H3K36
KMT3B	NSD1			H3K36
KMT3C	SYMD2			H3K36
KMT4	DOT1L		Dot1	H3K79
KMT5A	Pr-SET7/8	PR-set7		H4K20
KMT5B	SUV4-20H1	Suv4-20		H4K20
KMT5C	SUV4-20H2			
KMT6	EZH2	E(Z)		H3K27
KMT7	SET7/9			H3K4
KMT8	RIZ1			H3K9

Names of each modifying enzyme are indicated for several species (Human, Drosophila and S.cerevisiae) as well as the common nomenclature. Marks indicated in green and red have been associated with active and repressed transcription respectively (H = histone; K = Lysine; KAT = Lysine acetyl-transferase; KMT = Lysine methyl-transferase).
Adapted from Allis et al., 2007

Table 1.2 Deacetylation and demethylation of lysine residues

a

		Subcellular location	Knockout mice phenotype
Histone deacethylases			
Class I	HDAC1	nucleus	E10.5 proliferation defects
	HDAC2	nucleus	P1 cardiac malformation
	HDAC3	nucleus	E9.5 gastrulation defects
	HDAC8	nucleus	P1 craniofacial defects
Class IIa	HDAC4	nucleus/cytoplasm	chondrocyte differentiation defect
	HDAC5	nucleus/cytoplasm	cardiac hypertrophy after stress
	HDAC7	nucleus/cytoplasm	E11 endothelial dysfunction
	HDAC9	nucleus/cytoplasm	cardiac hypertrophy after stress
Class IIb	HDAC6	mainly cytoplasm	increased tubulin acetylation
	HDAC10	mainly cytoplasm	not known
Class III	SIRT1	nucleus	defects in spermatogenesis/germ cell function
	SIRT2	mainly cytoplasm	no obvious phenotype
	SIRT3	mitochondria	no obvious phenotype
	SIRT4	mitochondria	no obvious phenotype
	SIRT5	mitochondria	no obvious phenotype
	SIRT6	nucleus	premature aging
	SIRT7	nucleolus	reduce lifespan, cardiac hypertrophy

b

New nomenclature	Human	Drosophila	S. cerevisiae	Targeted residue
Histone demethylases				
KDM1	LSD1/BHC110	Su(var)3-3		H3K4me1/2; H3K9me1/2
KDM2			Jhd1	H3K36me1/2
KDM2A	JHDM1a/FBXL11			H3K36me1/2
KDM2B	JHDM1b/FBXL10			H3K36me1/2
KDM3A	JHDM2a			H3K9me1/2
KDM3B	JHDM2b			H3K9me
KDM4			Rph1	H3K9me2/3; H3K36me2/3
KDM4A	JMJD2A/JHDM3A			H3K9me2/3; H3K36me2/3
KDM4B	JMJD2B			H3K9me2/3; H3K36me2/3
KDM4C	JMJD2C/GASC1			H3K9me2/3; H3K36me2/3
KDM4D	JMJD2D			H3K9me2/3
KDM5		Lid	Jhd2	H3K4me2/3
KDM5A	JARID1A/RBP2			H3K4me2/3
KDM5B	JARID1B/PLU-1			H3K4me1/2/3
KDM5C	JARID1C/SMCX			H3K4me2/3
KDM5D	JARID1D/SMCY			H3K4me2/3
KDM6A	UTX			H3K27me2/3
KDM6B	JMJD3			H3K27me2/3

a. Based on sequence homology, HDACs were classified into four groups. The subcellular location of each enzyme as well as the phenotype of knockout mice are indicated.
Adapted from Mund and Lyko, 2010.
b. Marks indicated in green and red have been associated with active and repressed transcription respectively (KDM = Lysine demethylase).
Adapted from Allis et al., 2007.

Based on sequence similarity and cofactor dependency, 18 HDACs were divided into the “classical” and the “sirtuin” families (Haberland et al., 2009). Class I (HDACs1, 2, 3, 8), class IIa (HDACs 4, 5, 7, 9), class IIb (HDACs 6, 10) and class IV (HDAC11) belong to the classical family and require Zn^{2+} as an acetyl group acceptor. The class III enzymes (SIRT1-7) belong to the sirtuin family and need NAD+ for deacetylase activity (Yang and Seto, 2007). HDAC enzymes have little specificity for specific histones or lysine residues. However, establishment of knockout cell lines or mice have shown that the various isoforms are expressed in different tissues and at different times in development. For instance, HDAC4 plays a crucial role in skeletogenesis while HDAC 5 and 9 are important for cardiovascular development (Haberland et al., 2009).

The methylation state of the histone tails also depends on the antagonistic actions of large families of enzymes. Lysine residues are methylated by lysine methyl-transferases (KMTs) that contain a conserved SET-domain responsible for the transfer of a methyl group from S-adenosyl-L-methionine (SAM) leading to the production of S-adenosyl-L-homocysteine (SAH). SET-domain containing enzymes have been classified into seven main families (SUV39, SET1, SET2, EZ, RIZ, SMYD and SUV4-20) according to sequence homology and function. Unlike HATs, KMTs are residue-specific. For instance the SUV39 family of enzymes is responsible for H3K9 methylation while the SET1 family induce methylation of H3K4 residues (Dillon et al., 2005). Although arginine methylation has been less studied, enzymes implicated in the process were identified. Peptidylarginine methyltransferases (PRMTs) have been classified into four types, which are all able to monomethylate arginines. In addition, type I and type II PRMTs respectively catalyse asymmetric and

symmetric dimethylation of arginines. PRMTs were shown to target proteins involved in gene regulation including histones H3, H4 and H2A (Bedford, 2007).

Lysine methylation has been considered as a dynamic modification since the discovery of lysine demethylases (KDMs), although not all residues have had a demethylase identified so far. The first group of KDMs contains the flavin-dependent enzymes LSD1 that demethylates H3K4me1/2 (Shi et al., 2004) and H3K9me1/2 (Metzger et al., 2005) and LSD2 that is specific for H3K4me1/2 (Karytinos et al., 2009). A second family of KDMs is characterised by a Jumonji C (JmjC) domain and an oxygenase activity dependent on Fe(II) and α-ketoglutarate. Several members were shown to remove all three histone lysine-methylation states on H3K9 and H3K36 (Klose et al., 2006). Very little is known about arginine demethylation. The Jumonji domain-containing 6 protein (JMJD6) was described as a H3R2 and H4R3 demethylase (Chang et al., 2007) and PADI4 was shown to antagonise arginine methylation on H3 through deimination of residues (Cuthbert et al., 2004).

As previously mentioned, in addition to acetylation and methylation, histone proteins are subjects to a wide variety of post-translational modifications. Kinases and phosphatases respectively add and remove phosphate groups on serine, threonine and tyrosine residues while ubiquitin-ligases and de-ubiquitin enzymes control lysine ubiquitination levels on H2A and H2B (Bannister and Kouzarides, 2011).

1.3.1.2 Regulation of enzymes

Most modifying enzymes depend on cofactors that are products of the cellular metabolism. For instance, HATs transfer acetyl groups that come from acetyl-CoA. This compound is produced during the pyruvate decarboxylation reaction before entering the citric acid cycle. The availability of acetyl-CoA therefore depends on glucose metabolism, pyruvate being a product of glycolysis (Spriet and Heigenhauser, 2002). Histone methylases and DNA methyl-transferases on the other hand use SAM as a cofactor. Methionine synthetase, responsible for the conversion of homocysteine into methionine, uses methylated tetrahydrofolate (MTHF) as a source of methyl group. MTHF is a folate derivative, providing another link between gene expression and external signals (here dietary folate intake) (Loenen, 2006). Demethylation involves Fe2+ as a cofactor, showing again the dependence of enzymes on nutrient intake and the basic metabolism. In addition deacetylation was shown to be NAD+-dependent, whereas phosphorylation was shown to be ATP-dependent, both molecules being important cofactors in many metabolic reactions.

Enzymes that catalyse deposition and removal of histone modifications are very often found in conserved multi-subunit complexes where the associated proteins usually regulate their enzymatic activity, substrate recognition or cofactor binding. For instance, Class I HDACs assemble in at least four repressor complexes: Sin3, NuRD, CoREST and NCor/SMRT. The Sin3 complex contains among others HDAC1 or 2 along with the histone binding protein RbAp46/48, the histone demethylase KDM5A and SAP30 that mediates interaction with DNA-binding proteins. HDAC1 and 2 also interact with the methyl-CpG binding domain protein MBD3, the histone demethylase

LSD1 and the chromatin remodeler CHD4, a component of the NuRD complex. Other HDACs assemble complexes (e.g. HDAC3 with JMJD2 in the NCor/SMRT complex) or are able to bind directly to transcription factors (e.g. HDAC4 represses Myocyte enhancer factor 2 (MEF2)) (Youn et al., 2000).

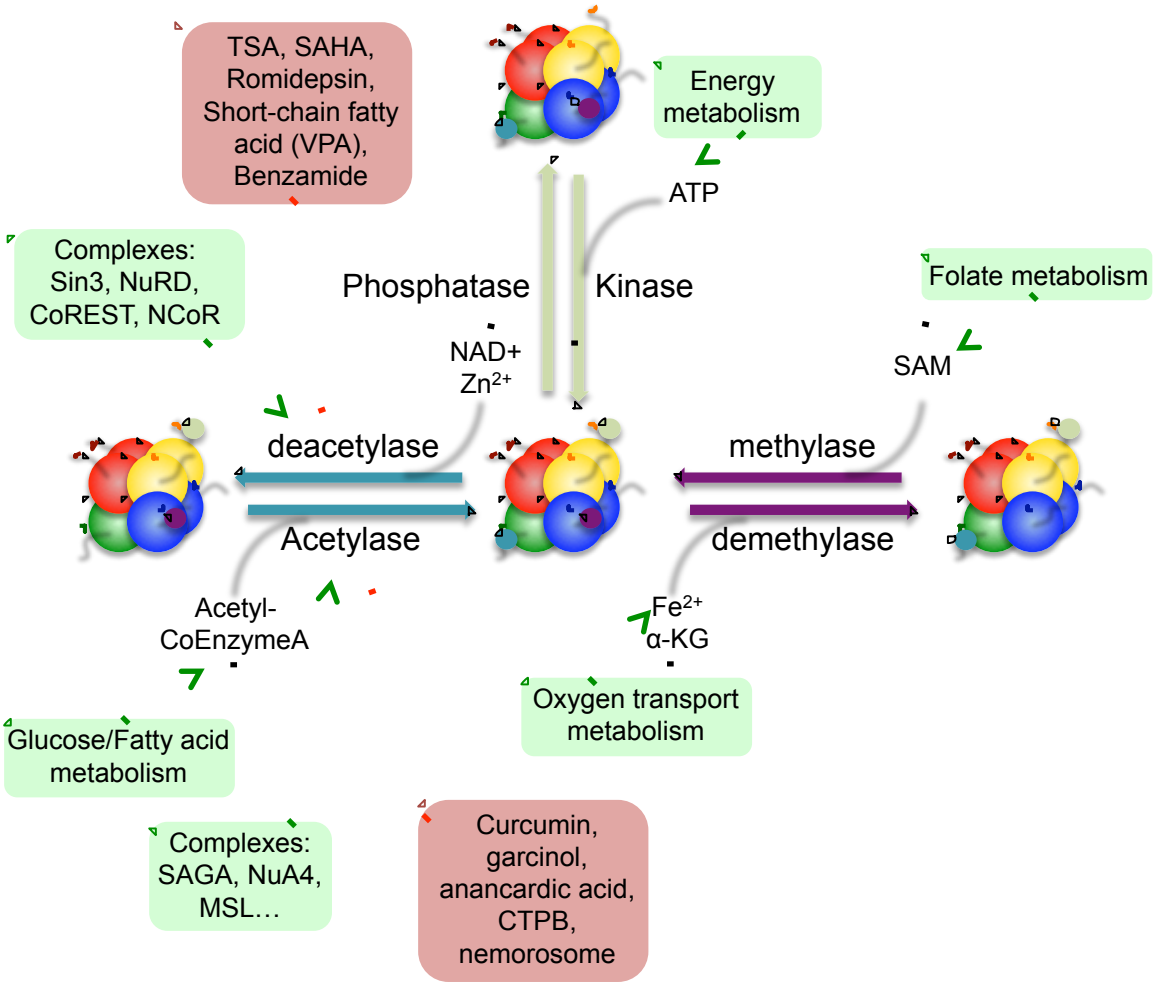
Similarly, HAT complexes have been identified and they mainly contain enzymes from the GNAT and MYST families. For instance in *Saccharomyces cerevisiae*, SAGA (PCAF in human) or SLIK complexes associate with the HAT Gcn5, whereas members of the MYST family are found in the NuA4 or SAS complexes (TIP60 in human and MSL in Drosophila). Those complexes play important roles in gene activation or DNA repair through interaction with histone modifications or histone variants (Lee and Workman, 2007).

Enzymes regulating acetylation levels have been implicated in several diseases. In fact, the identification of specific targets of HATs and HDACs revealed their role in cell proliferation, inflammatory response, learning and memory, neurodegeneration or viral infections. For this reason, small molecules that modulate HAT and HDAC activities have been extensively studied for potential therapy. A large number of HAT inhibitors were identified including the natural products curcumin, garcinol or anacardic acid (Dekker and Haisma, 2009). In addition, a few molecules were shown to activate HATs, probably through structural alteration (nemorosome, (Selvi et al., 2010). The study of epigenetic profiles in cancer cell lines revealed important alterations of DNA methylation or histone modification associated with alterations of gene expression including silencing of some tumour suppressor genes. As a result, several small molecules modulating activity of DNMTs or HDACs are currently being

tested in clinical trials or have been approved for cancer therapy (Mund and Lyko, 2010). The best-characterised DNMT inhibitor, 5-azacytidine was approved for the treatment of myeloid leukaemia. HDAC inhibitors have been classified into five different groups with different HDAC-specificities. For example, Trichostatin A (TSA) and Suberoylanilide hydroxamic acid (SAHA) belong to the hydroxamic family and inhibit all HDACs. Romidepsin belongs to the cyclic tetrapeptide family and inhibits HDAC1 and 2, whereas the short-chain fatty acids Valproic acid (VPA) and phenyl butyrate inhibit class I and IIa HDACs. Finally, HDAC1 and 3 are inhibited by the benzamide Etinostat and the synthetic benzamide N-acetyldinaline (Mund and Lyko, 2010) (**Figure 1.6**).

It is important to note that many mechanisms regulate modifying enzymes. For instance, the autoacetylation of HATs as well as post-translational modifications including phosphorylation or methylation. In addition, SIRT2 was shown to deacetylate p300, whereas HDAC3 targets PCAF (Selvi et al., 2010).

Figure 1.6 Regulation of modifying enzymes



Histone modifications are added and removed from histone tails by highly specialised families of enzymes. Importantly, those enzymes are regulated by many external signals including co-factors or small molecule inhibitors. In addition, the activity of modifying enzymes are dependant on several metabolic intermediates and therefore determined by the basic metabolism of the cell. Green boxes indicate activating events whereas red boxes contain repressive molecules.

1.3.1.3 Valproic acid

Valproic acid (VPA) is a branched short-chain fatty acid that was first synthesized in 1882. Used as a solvent for organic compounds until 1963, it was found to prevent pentylenetetrazol-induced convulsions in rodents (Meunier et al., 1963). Since then, this compound and its derivatives have been widely used as anti-epileptic drugs. The mechanism of VPA action remains uncertain, but for epilepsy, two main pathways have been identified. VPA is believed to alter the activity of the neurotransmitter GABA (Mesdjian et al., 1982) and to inhibit the activity of several ion channels (Tian and Alkadhi, 1994). VPA has been employed in many clinical trials for the treatment of various neurological disorders, but showed very little benefit, however it is an efficient drug for the treatment of several myeloid and lymphoid malignancies as well as some solid tumours (Chateauvieux et al., 2010).

In 2001, it was shown that VPA induces hyperacetylation of the N-terminal tail of histone H3 and H4. VPA was found to inhibit some Histone deacetylase (HDAC) enzymes, likely through binding to the catalytic site. Furthermore, it was shown to induce the differentiation and apoptosis of carcinoma cells in vitro and in vivo. Being well tolerated and very stable, VPA appears to be a promising drug for cancer treatment (Gottlicher et al., 2001). VPA inhibits Class I and class IIa HDACs. Interestingly, it increases the expression of HDAC9 and the class IV HDAC11 in human leukemic HL60 cells (Bradbury et al., 2005). HDAC expression is enhanced in many cancer types and in addition to targeting these proteins, VPA induces apoptosis and cell cycle arrest (Bolden et al., 2006).

1.3.2 Distribution of histone modifications

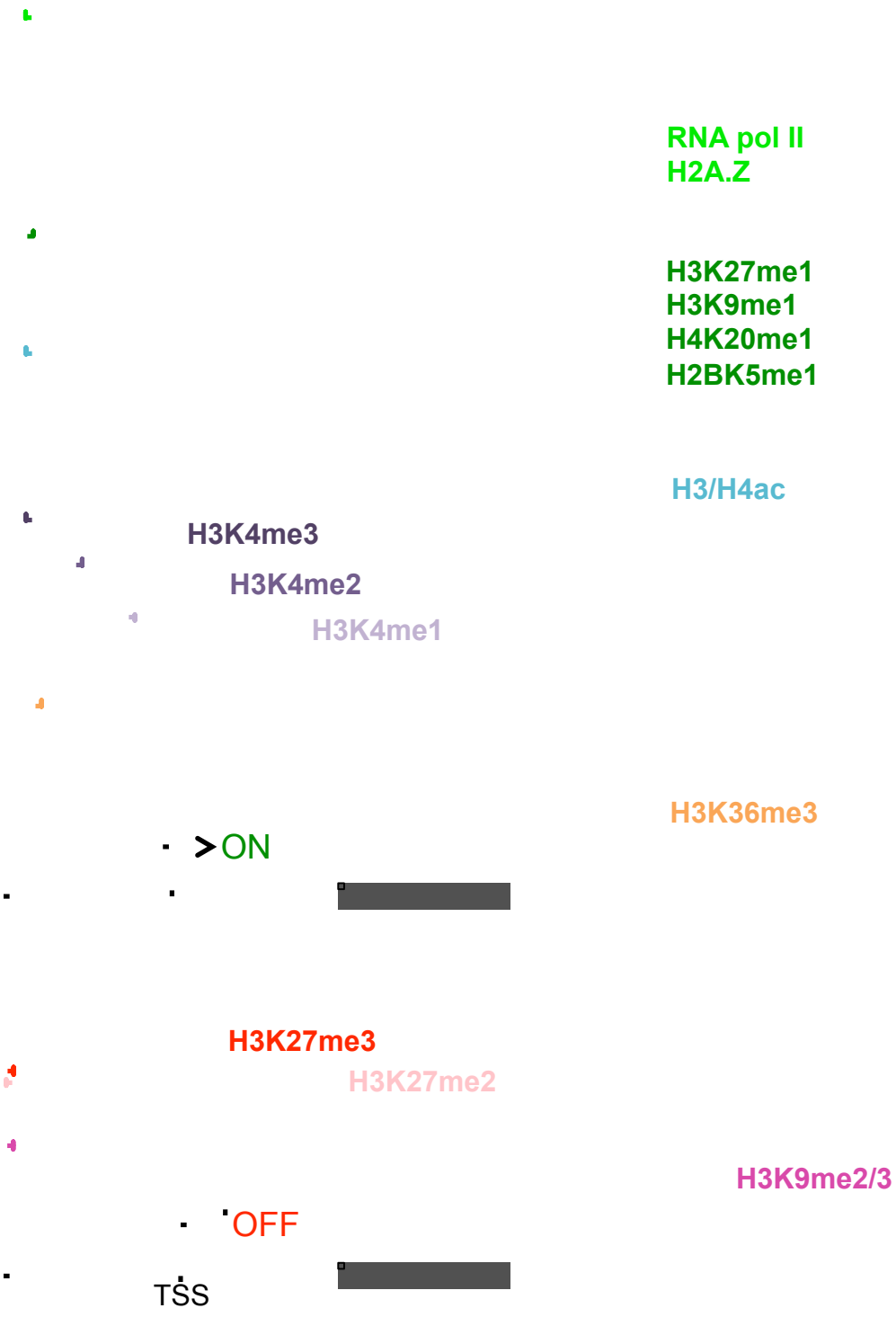
The use of chromatin immunoprecipitation (ChIP), coupled to high throughput technologies like microarrays (ChIP-chip) or next generation sequencing (ChIP-seq), has allowed the construction of maps of histone modifications on a “typical gene” (**Figure 1.7**). It is important to remember that the result of a ChIP depends on the optimization of the technique as well as the antibody specificity or how the chromatin is prepared (for instance fixed or unfixed). ChIP also generally involves an asynchronous population of cells and can not determine if two different modifications occur on the same histone, or the same nucleosome or even in the same cells. However, many studies showed consistent data and allowed correlations between the presence of specific histone marks and the transcriptional state of a gene (Lee and Mahadevan, 2009).

An early study on yeast using ChIP-chip revealed that nucleosome occupancy correlates negatively with transcription while H3K9ac, H3K14ac, H3K4me3 target promoters of active genes. In addition, H3K36me3 was detected in the bodies of active genes (Pokholok et al., 2005). Subsequent ChIP coupled with serial analysis of gene expression in primary human T cells confirmed that H3K9ac and H3K14ac co-localise with H3K4me3 at the promoter of constitutively active genes or T cell-specific genes, and that H3K27me3 was present on silent genes. Interestingly, H3K4me3 and H3K27me3 were sometimes found together on genes (Roh et al., 2006). In 2007, the first ChIP-seq allowed large scale mapping, at single nucleosome resolution, of 20 histone modifications as well as the histone variant H2A.Z, RNA polymerase II (RNA polII) and CTCF (Barski et al., 2007). Positive correlations with

transcription were observed for RNA pol II occupancy and H3K4me1/2/3. However, RNA pol II and H3K4me3 were also found on 37% and 54% of silent promoters respectively, indicating that those marks are not exclusive to active transcription. Interestingly, the methylation state of H3K4 was shown to decrease progressively from the transcription start site (TSS) into the gene body. In addition, positive correlations with transcription were established for H3K27me1, H3K9me1, H3K36me3, H3K20me1 and H2A.Z. In contrast, broad “blankets” of H3K27me3 were detected over silent genes as well as peaks of H3K9me2/3. Finally, CTCF was shown to separate regions of different chromatin states (Barski et al., 2007).

In embryonic stem (ES) cells, the active mark H3K4me3 and the repressive mark H3K27me3 were found to co-localise on certain promoters called “bivalent domains”, and are believed to allow a “poised” state for subsequent activation or silencing (Bernstein et al., 2006; Mikkelsen et al., 2007). In another study in *C. elegans* using ChIP-chip, H3K36me3 was showed to localise to the body of transcribed genes. More precisely, the enrichment was found in exonic regions in opposition to intronic regions, and has been confirmed in mouse and human (Kolasinska-Zwierz et al., 2009). In addition, enhancers were shown to be highly enriched in H3K4me1, H2A.Z and H3K27ac. However, analysis of ENCODE (Encyclopedia of DNA Elements) data showed that the chromatin state of enhancers is highly cell specific compared with promoters or insulators (Barski et al., 2007; Heintzman et al., 2009).

Figure 1.7 Distribution of histone modifications



The use of chromatin immunoprecipitation coupled with microarrays (ChIP-chip) or sequencing (ChIP-seq) allowed the construction of maps of typical histone modifications on active and inactive genes.
Adapted from Barski et al., 2007.

1.3.3 Effect of histone modifications on chromatin structure and transcription

1.3.3.1 Charge neutralisation

The best-characterised example of chromatin modulation due to a change of charge is histone acetylation. This neutralises the positive charge of lysines and therefore reduces the association of DNA with the histone tails, thereby increasing the DNA accessibility to non-histone proteins. This is consistent with reconstitution of nucleosome arrays which showed that high levels of acetylation inhibited higher order folding of the chromatin, and increased transcription (Tse et al., 1998). In addition, arrays containing recombinant histone H4 homogeneously acetylated at lysine 16 could not fully condense into the 30mm fibre in vitro (Robinson et al., 2008; Shogren-Knaak et al., 2006), consistent with the idea that H4K16 has a unique impact on chromatin folding (Dorigo et al., 2003). In addition, microarray analyses performed after histone H4 was mutated at lysine 5, 8, 12 and 16 in yeast, found that the global changes in transcription induced by mutation of K16 did not correlate with the transcriptional outcomes of other mutations, reinforcing the special role of H4K16ac (Dion et al., 2005).

1.3.3.2 Binding histone modifications

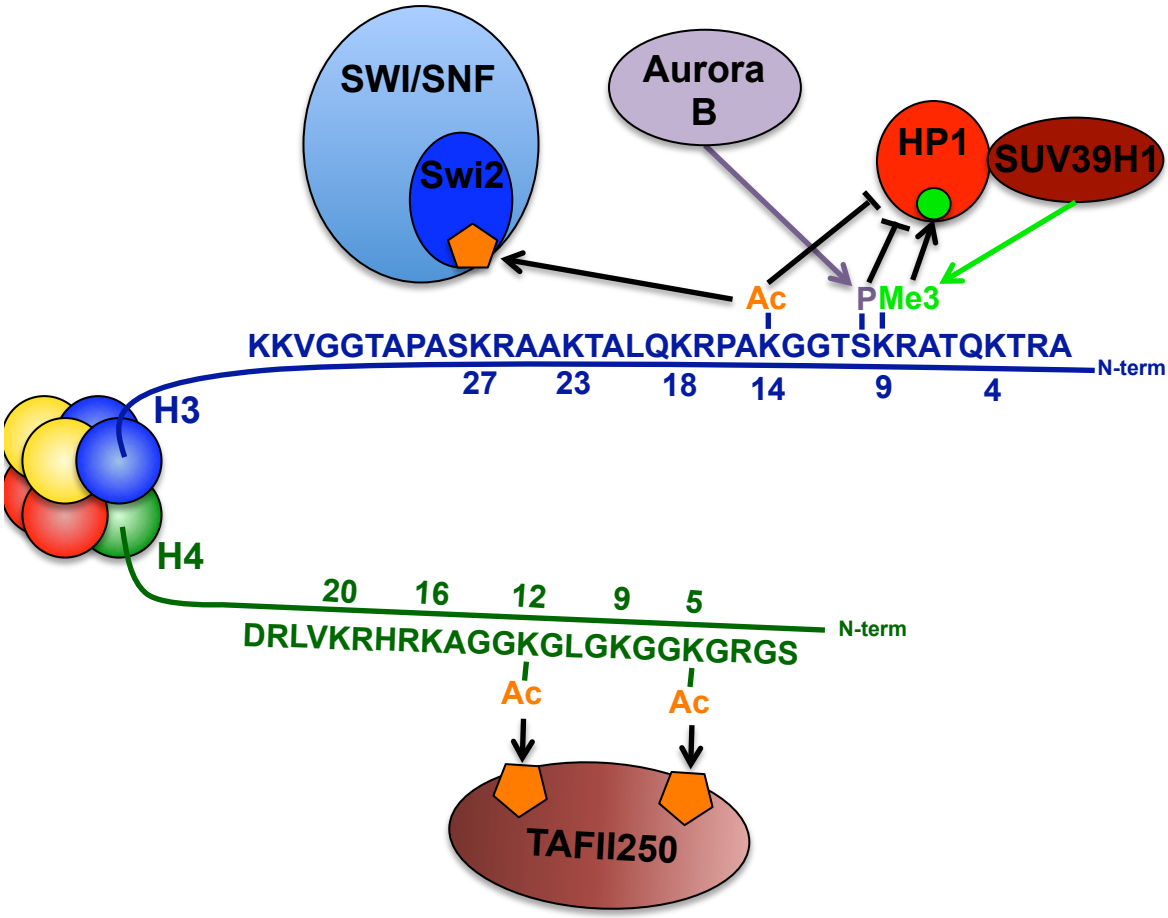
Many chromatin-associated proteins possess particular domains able to recognise and bind specific modifications. For instance, bromodomains are capable of interaction with acetyl residues while methyl marks are recognised by a wide range of domains including chromodomains, Tudor domains or PHD fingers. The family of 14-3-3 proteins has been implicated in phosphoserine-binding. Along with the diversity of histone modifications, the diversity of "reader" domains provide a new explanation for the multiple outcomes of chromatin marks (Taverna et al., 2007). In fact, reader domains give a chance for modifying enzymes and chromatin-binding proteins to communicate. Thereby, a series of modifications can act together to regulate transcription in a specific manner (Jenuwein and Allis, 2001; Spotswood and Turner, 2002).

All chromatin remodelling complexes, which move nucleosomes to regulate DNA accessibility, recognize histone modifications and interact with chromatin modifiers (Clapier and Cairns, 2009). For instance, the member of the SWI/SNF complex swi2 possesses a bromodomain required for acetylated nucleosome displacement (Chandy et al., 2006). In addition, the initiation complex component TAFII250 contains a bromodomain that specifically binds to acetylated H4 on lysines 5 and 12, suggesting a crucial role of acetylation in transcriptional initiation (Jacobson et al., 2000). Finally, the ATP-remodeler CHD1 that belongs to the HAT SAGA complex contains two N-terminal chromodomains that recognise methylated H3K4. This complexity supports the idea of cross-talk between different histone marks and a code for transcription (Pray-Grant et al., 2005).

The formation of heterochromatin for gene silencing or chromosome segregation also necessitates the recognition of specific histone modifications. The heterochromatin protein HP1 is a fundamental factor in chromatin packaging and possesses a N-terminal chromodomain that specifically binds to H3K9me2/me3 (Bannister et al., 2001). Biochemical analysis indicates that HP1 interacts with SUV39H1, the enzyme responsible for H3K9 methylation. However G9a, another enzyme inducing H3K9 methylation, did not recruit HP1, suggesting a silencing mechanism independent of HP1 (Stewart et al., 2005). Interestingly, the histone code hypothesis was well illustrated by the description of a “methyl-phospho” switch. Methylation often occurs in ARKS motifs where a serine residue adjacent to a modified lysine can be phosphorylated. In fact, the phosphorylation of H3S10 by Aurora B type kinases was shown to block HP1 binding during M-phase of the cell cycle (Fischle et al., 2005; Fischle et al., 2003). However, a more complex picture appeared when the role of H3K14ac was described in HP1 eviction (Cheung et al., 2000) whereas H3S10p and H3K9me were shown to co-localise and not prevent HP1 binding. These studies showed that HP1 could bind to all but H3K9me-S10p-K14ac peptides (Mateescu et al., 2004) (**Figure 1.8**).

All these studies strongly suggested a role for histone modifications as binding sites in transcription activation or heterochromatin formation. In addition, the grouping of similar proteins that recognise different histone marks supports the hypothesis of a histone code.

Figure 1.8 Crosstalk between histone-binding proteins



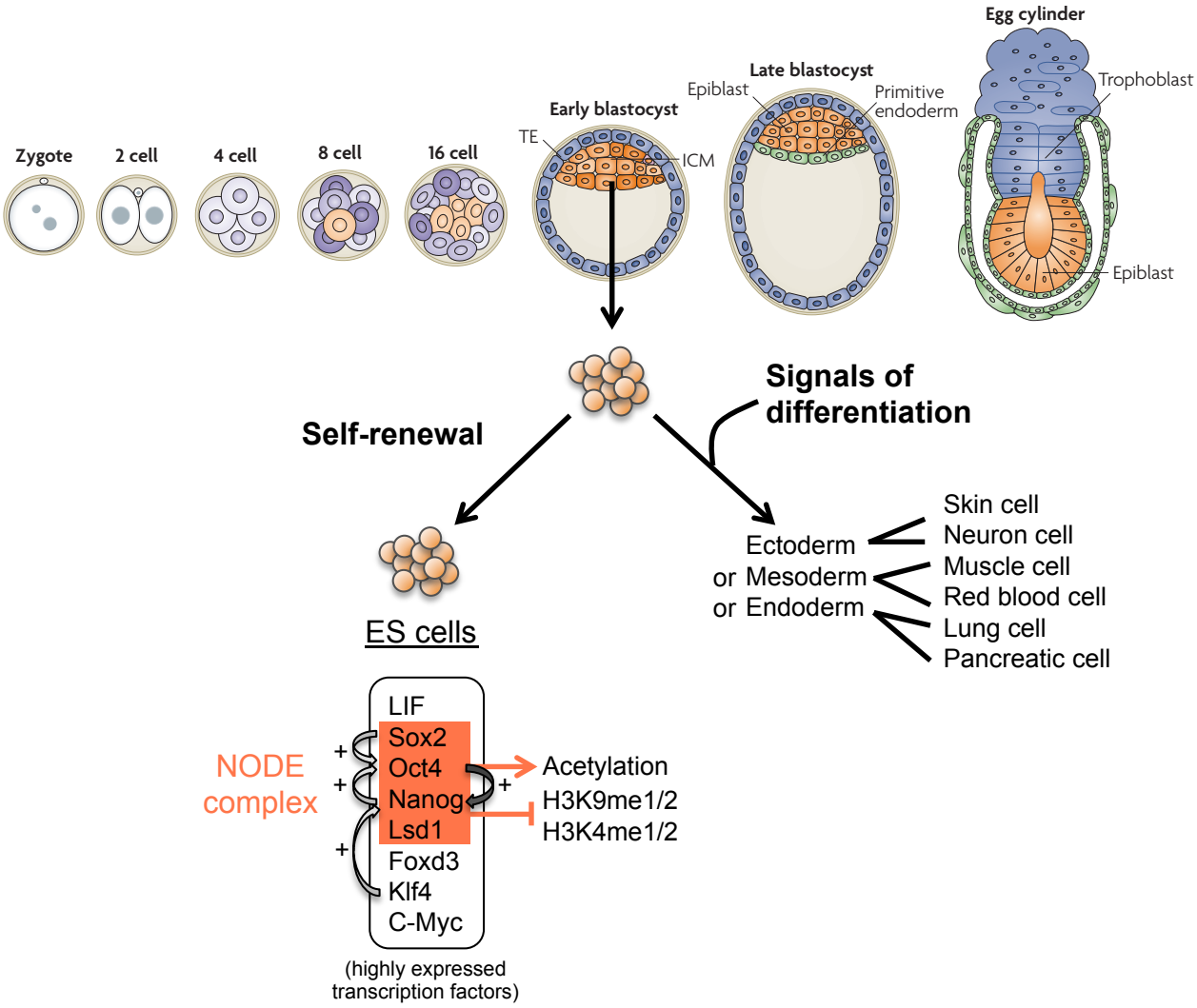
Histone modifications are recognised by specific motifs of binding proteins. Different proteins bind to different modifications. Therefore, when protein readers interact with each others, they lead to cross-talk between histone marks. For instance, the complex SWI/SNF recognises H3K14ac and, using ATP energy, displaces nucleosomes to favour transcription. In addition, TAFII250 lead to the assembly of RNA polII machinery when bound to H4K12,K5ac. H3K14ac together with H3S10P inhibit HP1 binding to H3K9me3, implicated in heterochromatin formation. HP1 interact with the SUV39H1 methylase and recognises H3K9me2/3.

1.4 Embryonic stem cells

In 1981, two groups independently established the first mouse ES cell lines derived from the inner cell mass of late blastocysts (Evans and Kaufman, 1981; Martin, 1981). Since then, gene targeting experiment in ES cells have become a major tool for the study of gene function and gene mutation in living mice (Capecchi, 2005). Embryonic stem (ES) cells are characterised by two main features: the capacity for prolonged self-renewal and the potential to differentiate into almost any cell type. In fact, ES cells injected into tetraploid blastocysts (required for extra-embryonic tissue formation) give rise to an entire embryo, confirming their pluripotent nature. This also shows that ES cells are pluripotent and not totipotent as they need the blastocyst structure to generate a chimeric animal (i.e. They can't generate extra-embryonic tissues) (Nagy et al., 1993). The pluripotent nature of ES cells, can now be reproduced in culture due to many early studies monitoring ES cell differentiation pathways and identifying specific markers. Based on those studies, protocols that involve the spatial rearrangement of cells in culture and exposure to specific factors were described in order to drive differentiation of ES cells down defined lineages (**Figure 1.9**) (Murry and Keller, 2008).

For many reasons ES cells constitute an invaluable tool which could potentially lead to the understanding of mechanisms governing pluripotency and differentiation. Differentiation of ES cells in culture offers the possibility to study the establishment and long-term maintenance of specific gene expression programmes that define cell identity. A full understanding of pluripotency and cell fate is crucial for the development of regenerative medicine. Also, ES cells have allowed better

Figure 1.9 Embryonic stem cells



Embryonic stem (ES) cells derived from the inner cell mass (ICM) of early blastocyst. It is possible to maintain them in a pluripotent state with the addition of LIF in the culture media. A cocktail of transcription factors are highly expressed in ES cells. They regulate each other and many common targets for pluripotency maintenance. Also, they interact with chromatin modifying enzymes including the demethylase LSD1 in a complex named NODE.
ES cells can be differentiated in culture. Using various signals, ES cells can be driven down many differentiation pathways.
Frame of mouse embryonic development taken from Hemberger et al., 2009.

understanding of tumour formation as cancer cells often present a stem cell-like phenotype. Indeed, the first embryonic cell lines were derived from mouse teratocarcinomas (Solter, 2006), and embryonal carcinoma cells allowed the identification of alkaline phosphatase activity as a marker of undifferentiated cells (Berstine et al., 1973) or the increase of SSEA-1 expression at an early stage of differentiation (Solter and Knowles, 1978).

1.4.1 Genetic regulation of pluripotency

In vivo, pluripotency is a transitional state, and very early in development most of the cells that compose an embryo are engaged in a differentiation pathway. However during this short period of time, cells have to maintain the pluripotent state and a high rate of proliferation. Interestingly, many studies have shown that gene expression in ES cells is very variable and "noisy", leading to heterogeneous populations poised for a range of lineage specification. It is therefore difficult to define what makes a cell pluripotent (Silva and Smith, 2008).

Nevertheless, gene expression networks have been described in ES cells. In 2006, it became possible to generate induced pluripotent stem (IPS) cells. Transfection in mouse fibroblasts led to the conclusion that only four transcription factors were required for reprogramming: *Sox2*, *Oct3/4*, *Klf4* and *c-Myc*. IPS cells showed high similarity to ES cells and were able to differentiate into mouse embryos when injected into blastocysts (Takahashi and Yamanaka, 2006). A well-known signalling pathway in ES cells is initiated by the cytokine Leukaemia inhibitory factor (LIF) that acts via the glycoprotein-130 receptor to promote a JAK-STAT3 signalling cascade playing a

crucial role in maintenance of undifferentiated state (Raz et al., 1999). In culture, LIF is sufficient to maintain pluripotency of ES cells but also proliferation by targeting the transcription factor c-Myc (Cartwright et al., 2005).

Deletion of *Sox2* induced ES cells to differentiate, showing that *Sox2* is indispensable for maintaining pluripotency. *Sox2* and *Oct4* bind together on enhancers linked to pluripotency associated genes like *lefty1* or *fgf4*. The elimination of *Sox2* did not affect those genes, possibly because of redundant members of the Sox family. However, *Oct4* was down-regulated in *Sox2*-null ES cells, suggesting a role for *Sox2* in maintaining *Oct4* expression (Masui et al., 2007). In fact, the transcription factor *Oct4* has been extensively studied due to its major role in pluripotency maintenance. *Oct4* overexpression initiates differentiation into primitive endoderm whereas repression leads to the trophectodermal lineage (Niwa et al., 2000). Finally, *Nanog* was identified as another major player in pluripotency maintenance. When constitutively expressed in ES cells, it is able to maintain pluripotency without LIF. In addition, loss of pluripotency and extraembryonic ectodermal differentiation was observed in *Nanog*-null ES cells (Mitsui et al., 2003).

Although a pool of transcription factors was shown to be indispensable for pluripotency maintenance in ES cells, how they are regulated and how they regulate their targets is still under investigation. As mentioned earlier, Sox2 is a positive regulator of *Oct4* (Masui et al., 2007). In addition, together with FoxD3, Nanog was found to be another enhancer of Oct4 expression, and is repressed by Oct4 itself. The suggested negative feedback loop was confirmed when overexpression of *Nanog* or *FoxD3* did not change *Oct4* expression. On the other hand, Oct4 was

shown to bind *Nanog* promoter for a dose-dependant regulation. In fact, Oct4-siRNA initially raised *Nanog* expression but an increase in siRNA transfection led to *Nanog* downregulation (Pan et al., 2006). As a final example, ChIP-chip experiments gave evidence that Klf4, Klf2 and Klf5 bind to several pluripotency-associated genes and an enhancer that interacts with those factors was found near Nanog promoter (Jiang et al., 2008).

Several high-throughput studies have tried to identify targets of the master regulators of pluripotency. Chromatin immuno-precipitation (ChIP) is a technique by which specific antibodies separate bulk chromatin according to the presence or not of specific histone modifications or non-histone proteins. Bound and unbound (or input) fractions of chromatin are then analysed by qPCR, microarrays or sequencing in order to measure enrichment. ChIP combined with DNA microarray experiments performed using human ES cells, revealed that Oct4, Sox2 and Nanog share a significant number of targets. The three transcription factors were shown to co-localise on 353 promoters of protein-coding genes and to actively regulate those targets. In fact, half of the co-occupied genes were transcriptionally active including some participating in pluripotency maintenance while the other half corresponded to inactive genes, mostly transcription factors involved in differentiation pathways (Boyer et al., 2005).

1.4.2 Chromatin structure in ES cells

Because it is now well accepted that chromatin structures are associated with the transcriptional activity of genes, the purpose of many recent studies has been to describe a putative epigenetic signature of ES cells. The challenge for those cells is to maintain pluripotency while preparing for differentiation, to repress developmental genes but keep them poised for transcription.

Interestingly, the reprogramming of primary human fibroblasts was shown to occur via the injection of only two transcription factors (Oct4 and Sox2) complemented with Valproate (VPA) treatment. The IPS cells then generated were able to regain an ES cell-like DNA methylation profile and global expression pattern and to differentiate into the three germ lineages (Huangfu et al., 2008). The fact that the HDAC inhibitor VPA was able to increase the efficiency of reprogramming supports the idea that chromatin structure is an important specifier of ES cell identity. Indeed, chromatin in ES cells was described as being less compact and acetylation on histones H3 and H4 have been shown to decrease during differentiation. ES cells seem to present a more diffuse heterochromatin structure as HP1 and H3K9me3 immunostaining appeared in large and poorly defined regions. However the clustering of these heterochromatin components in visible foci was observed during differentiation (Meshorer et al., 2006). In addition, FRAP (fluorescence recovery after photobleaching) experiments suggest that the exchange rate of HP1, H1, H2B and H3 was higher in undifferentiated cells compared to differentiated ES cells, showing that those proteins are loosely bound to ES cell chromatin, supporting the idea of an open, accessible chromatin structure in ES cells (Meshorer et al., 2006).

Early replication during the S-phase of the cell cycle has been associated with gene rich transcriptionally active regions with high levels of histone acetylation. Following BrdU incorporation, it was observed that the number of genes replicating late increased during ES cell differentiation. Interestingly, early replication in ES cells or at a later stage during differentiation did not always correlate with transcription activity but did reflect lineage potential (Azuara et al., 2006).

Supporting the idea of a prepared chromatin structure, the most unusual chromatin signature in ES cells is the existence of “bivalent domains” (Bernstein et al., 2006). In this study, highly conserved non-coding element-rich regions enriched for important developmental transcription factors were analysed. It was often observed that large blankets of the repressive mark H3K27me3 co-localised with peaks of the activating mark H3K4me3. Bivalent domains were often associated with low expression, reinforcing again the concept of poised chromatin. It was proposed that, with both marks on the same promoter, a gene can be either switched on or off according to differentiation programme (Bernstein et al., 2006). CpG-rich promoters are generally depleted in DNA methylation and enriched for the H3K4me3 mark, however, genes associated with high-CpG content promoters (HCP) are not always transcriptionally active. It was shown in ES cells that a significant number of HCPs were also associated with H3K27me3 and were therefore defined as bivalent domains. Bivalent HCPs were often associated with silenced markers of differentiation while HPCs only containing H3K4me3 were associated with housekeeping genes. In addition, ChIP experiments examining histone marks in ES cells, neural progenitor cells and mouse embryonic fibroblasts, showed that during differentiation bivalency remains or

resolves in the erasure of one of the marks according to lineage commitment and gene activity (Mikkelsen et al., 2007).

A poised RNA polymerase II is characterised by phosphorylation on serine 5 (Ser5P) while an active polymerase is phosphorylated on serine 2 (Ser2P). Going again with the idea that bivalent genes are poised for differentiation, it was shown in ES cells that bivalent domains were associated with Ser5P and not with Ser2P (Stock et al., 2007). In addition to H3K27me3 marks, the Polycomb complex deposits monoubiquitination on lysine 119 of H2A (H2AK119ub1) through its component Ring1b. This mark was shown to retain the poised polymerase on bivalent domains and loss of Ring1b leads to the derepression of bivalent genes in ES cells (Stock et al., 2007). In an additional study using human ES cells, ChIP-chip screening of the entire genome showed that 74% of annotated promoters were enriched for H3K4me3 as well as H3K9ac, H3K14ac and the initiation form of RNA polymerase II (Guenther et al., 2007). However, enrichment of H3K36me3, a mark linked to transcription elongation, was only found on genes associated with detectable amounts of transcript. This suggests most genes in ES cells are poised for transcription (Guenther et al., 2007).

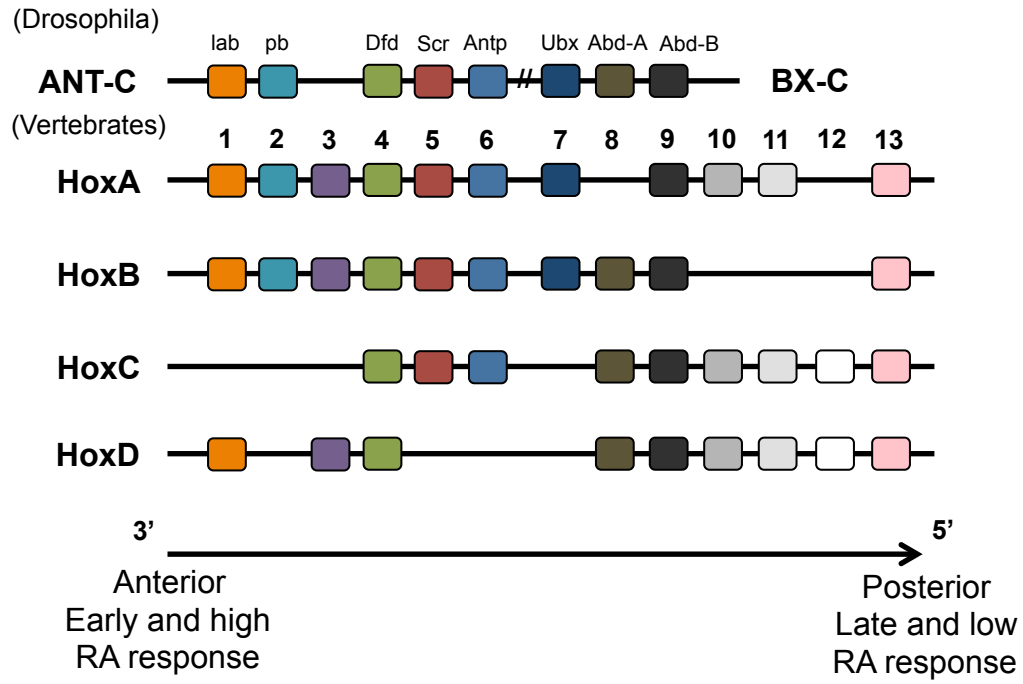
Although many genes are poised for transcription in ES cells, they have to remain silenced for maintenance of pluripotency. Indeed, the Polycomb mark H3K27me3 was shown to participate in the silencing of bivalent genes. An additional mark, H3K9me3, originally associated with heterochromatin, was also shown to be important for gene repression in ES cells and is observed on a large set of active and inactive genes, including developmental regulators, and co-localises with H3K4me3 and H3K27me3 (Bilodeau et al., 2009).

Many studies looking at various knockouts of, for instance, isoforms of DNA methyltransferases (dnmts), Polycomb components or some histone modifiers, showed that histone modifications are crucial for development but not indispensable for ES cell phenotype maintenance. From those observations came the conclusion that epigenetic modifications are not necessary for pluripotency. However, it is possible that the knockout of one component is not enough to disturb the entire pluripotency network (Meissner, 2010). In addition one could argue that although self-renewal was not affected in those knockout cell lines or embryos, differentiation was not possible and pluripotency in its strictest sense was indeed disturbed by loss of specific marks.

1.4.3 Homeobox (Hox) genes

Mutations in *Drosophila* leading to phenotypic changes, like the appearance of extra legs or extra wings, led to the discovery of a gene cluster that controls thoracic and abdominal segments. In fact, development of segment-specific body structures involves the regulation of two complexes of homeobox genes (hox genes) named Bithorax and Antennapedia. Mutations in specific genes were shown to induce ectopic expression of other members in the cluster and abnormal phenotype (homeotic transformations) (Frischer et al., 1986). Hox genes are highly conserved transcription factors containing a homeodomain that bind to enhancers. In vertebrates, 39 hox genes are divided into four clusters A, B, C and D located on chromosomes 6, 11, 15 and 2 in mice, probably resulting from a double duplication during evolution. One major feature of the hox genes is the correlation between their

Figure 1.10 The Hoxb cluster of genes



The developmentally regulated hox genes are highly conserved between species. They are found in clusters where their physical order mimics their temporal and spatial expression pattern during embryonic development. In vertebrates, a double duplication during evolution led to the establishment of four clusters containing 13 homologues. In *Drosophila*, two clusters of homeotic genes were identified. The Antennapedia complex (ANT-C) contains the genes labial (lab), proboscipedia (pb), Deformed (Dfd), Sex combs reduced (Scr), and Antennapedia (Antp). In the Bithorax complex (BX-C) are found the genes Ultrabithorax (Ubx), abdominal-A (abd-A), and Abdominal-B (Abd-B). 3' genes are expressed early during development and in the anterior part of the embryo while 5' genes are expressed later and in the posterior part of the embryo. Paralogues are indicated by the same colour.

physical order along the chromosome and their expression along the antero-posterior axis. In addition to the so-called “spatial colinearity”, a “temporal colinearity” of hox gene expression was observed as the anterior 3’ hox genes are expressed early during the development and the posterior 5’ hox genes are expressed much later (Krumlauf, 1994) (**Figure 1.10**). Homeotic transformations confirmed the important role of hox genes in development. Interestingly, it was recently shown that this set of genes is also expressed in the adult, in order to provide a certain level of plasticity, as differentiation also occurs in adult organisms (Daftary and Taylor, 2006).

1.4.3.1 Colinearity

The fact that hox genes are found in clusters suggests common mechanisms of regulation. Studies involving loss- or gain-of-function of several hox genes gave some evidence for the dominance of posterior genes. In fact, when a hox gene is mutated, the affected segment usually develops characteristics of the segment immediately anterior to it. However in vertebrates, mutations generally lead to less dramatic effects, probably due to compensation between paralogues. Those observations led to the “hox code” hypothesis, proposing that morphological identities in vertebrates result from the activity of a unique set of hox genes (Kessel and Gruss, 1991). For instance, the identity of the hindbrain rhombomere 4 (r4) is specified through auto- and cross-regulatory loops between hoxb1, hoxb2, hoxa1 and hoxa2. Using sequence alignment and targeted deletion, r4 restricted enhancers were identified near hoxa2 and hoxb2 (Tumpel et al., 2007).

In flies, the asymmetric distribution of hox genes allows patterning of the body. However, how temporal colinearity is achieved in vertebrates is still under debate and led to the “hox clock” hypothesis, which proposes several mechanisms. Timing of hox genes expression could be due to (1) passive or active chromatin remodelling along with progressive opening of the clusters, (2) changes in affinity of an enhancer for an effector molecule or (3) the action of regulatory sequence outside of the cluster (Kmita and Duboule, 2003).

A multitude of studies involving LacZ-transgene insertions in ES cells along with implantation in blastocysts allowed the monitoring of expression of hox genes during development. The purpose of those experiments was to show the regulatory role of clustering and therefore to show that spatial distribution of hox genes on the genome is crucial for temporal and spatial colinearity. For example, when hoxd11 and hoxd9 were inserted 5’ of the hoxd cluster, they showed both late ectopic expression and endogenous-like expression (van der Hoeven et al., 1996). Similar observations were made when hoxb1 and associated regulatory sequences were inserted at the end of the hoxd cluster. In addition the authors noticed an early induction of hoxd13 (Kmita et al., 2000) likely due to premature chromatin decondensation (Morey et al., 2008). Unfortunately, those experiments did not entirely clarify the temporal colinearity of hox genes as regulation appeared to depend on a combination of events including spatial distribution but also cis-regulators. However, it is likely that silencing mechanisms are in place to keep 5’ end genes silent suggesting progressive opening of the chromatin during development. In fact, when a specific portion at the end of the hoxd cluster was deleted, hoxd genes like hoxd10 and hoxd4 became early expressed (Kondo and Duboule, 1999).

As mentioned earlier, hox genes are able to respond to endocrine signals including retinoic acid (RA) exposure. In fact retinoic acid response elements (RAREs) were identified in the hoxa1, hoxb1, hoxa4, hoxb4 and hoxd4 loci (Barber and Rastegar, 2010). Supporting the idea of a progressive open environment in ES cells exposed to RA, the promoters of hoxb1 and hoxb9 were successively acetylated along with and prior to their induction. In addition, fluorescence in situ hybridization (FISH) experiments showed a progressive looping movement of the hoxb cluster outside the chromosome territory upon induction of transcription by RA (Chambeyron and Bickmore, 2004). Another study examined the progressive opening of chromatin at the hoxd4 locus following RA exposure (Rastegar et al., 2004). ChIP experiments on embryonal carcinoma cells and mouse embryos were used to look at active histone modifications, RNA polymerase II and CBP binding on different regions of the hoxd4 locus during gene induction. It was clear that active mark enrichment as well as CBP binding occurred first in the 3' enhancer and then on the promoter along with RNA pol II binding, supporting the opening chromatin model (Rastegar et al., 2004).

Finally, in addition to RARE-associated enhancers, several regulatory regions are known to influence hox gene expression. For instance, the CR3 enhancer common to hoxb3 and hoxb4 induces overlapping expression of those genes in r6 and r7, and is also implicated in hoxd4 regulation (Gould et al., 1997). Similarly, a cluster of enhancers covering 40kb identified upstream of the hoxd cluster was shown to control the expression of several unrelated genes including hoxd genes, Evx2 and lunapark (Spitz et al., 2003).

1.4.3.2 Polycomb/Trithorax regulation of hox genes

As previously mentioned, bivalent domains are present in ES cells, possibly allowing a gene to be both silent and poised for activation. A bivalent gene is characterised by the presence of both activating (H3K4me3) and repressing (H3K27me3) marks, deposited by the Trithorax (Trx) and the Polycomb (PcG) protein complexes respectively (Bernstein et al., 2006).

Polycomb proteins were originally discovered in *Drosophila* as negative regulators of Hox genes during segmentation (Lewis, 1978). Since then, an abundant literature has described a number of Polycomb complexes, as well as the antagonistic Trithorax group of proteins (Beisel and Paro, 2011; Kerppola, 2009; Margueron and Reinberg, 2011; Ringrose and Paro, 2004; Schuettengruber and Cavalli, 2009; Schuettengruber et al., 2007; Simon and Kingston, 2009; Simon and Tamkun, 2002). Polycomb proteins associate in two main complexes: Polycomb repressive complexes 1 and 2 (PRC1 and PRC2). The PRC2-type complex contains four core subunits: E(z), Esc/Escl, Su(z)12 and Nurf-55 in *Drosophila*. The signature activity of PRC2 is provided by E(z) of which the SET domain catalyzes di- and tri-methylation of H3K27 and H1K26. Esc and Su(z)12 are required for E(z) methyl-transferase activity and Nurf-55 is a histone binding protein. Many mammalian homologues have been found, providing a diversity of function for PRC2. For instance, the Esc homologue EED generates four isoforms from alternative translation initiation sites. The main PRC2 complex contains EED1 while the alternative complex PRC3 contains EED3/4 that methylates H3K27 only in the absence of H1 (Kuzmichev et al., 2004). EED2 is found in PRC4, active in undifferentiated cells. PRC4 preferentially

methylates H1 and contains the SIRT1 H1K26 deacethylase (Kuzmichev et al., 2004). In addition, there are two mammalian E(z) homologues, EZH1 and EZH2. Immunoprecipitation and ChIP-chip confirmed that both isoforms associate with the other core components and target the same set of genes. However, EZH1 showed less methyltransferase activity and fewer binding sites. Furthermore, EZH2 was found in proliferative tissues while EZH1 was present in differentiated tissues where it inhibits transcription and compacts the chromatin fibre (Margueron et al., 2008). Finally in *Drosophila*, Pcl was co-purified with PRC2 core components and Pcl mutants show a decrease in H3K27me3 levels and the derepression of target genes (Nekrasov et al., 2007). The mammalian homologue PHF1 was also shown to be required for the production of high H3K27me3 levels (Sarma et al., 2008). In *Drosophila*, PRC1 is composed of Pc, dRing, Psc and Ph. Importantly, Pc possess a chromodomain that recognises H3K27me3 and dRing is an E3 ubiquitin ligase that induces mono-ubiquitylation of H2AK119. Again, many homologues were characterised in mammals. In 2006, the polycomb protein Pho was identified in two complexes: PhoRC and Pho-dINO80. PhoRC (Pho-repressive complex) is composed of Pho, Phol and dSfmbt. Loss of hox gene silencing was observed in Pho/Phol double mutants as well as in dSfmbt mutants. ChIP experiments confirmed the co-localisation of the three components and the MBT (malignant brain tumor) repeats of dSfmbt are able to bind H3K9me1/2 and H4K20me1/2 (Klymenko et al., 2006). Two years later the dRAF (dRing-associated factor) complex was characterised. When dRing was immunoprecipitated from Pc-depleted extracts, Psc and some additional proteins were copurified including Raf1, the fly homologue of KDM2 that specifically demethylates H3K36me2. KDM2 enhances dRing activity but also Pc activity,

suggesting a cooperation between PRC1 and dRAF (Lagarou et al., 2008). Finally, the PR-DUB (Polycomb repressive deubiquitinase) complex was recently discovered, composed of the ubiquitin hydrolase Calypso and Asx and is required for long-term repression of hox genes. Calypso specifically targets H2A and derepression of the hox gene Ubx was observed in Calypso knockouts (Scheuermann et al., 2010) (**Table 1.3**).

Similarly, Trithorax complexes are composed of a heterogeneous panel of proteins. Importantly, the two Trithorax complexes TAC1 and MLL mediate the tri-methylation of H3K4, a mark associated with active transcription. The Trithorax complex BRM, related to the SWI/SNF complex, induces ATP-dependant chromatin remodelling. Finally, the ASH1 complex is responsible for H3K36 di-methylation, a mark associated with transcription elongation (Schuettengruber et al., 2007) (**Table 1.4**).

How Trithorax and Polycomb proteins are targeted to specific loci is still under debate. In *Drosophila*, the DNA sequence itself has been implicated with the discovery of Polycomb/Trithorax response element (PRE/TRE) sequences (Ringrose and Paro, 2007). PRE/TRE are composed of various binding motifs, including Pho, Zeste, Gaf, dsp1, Grh or Sp1 binding sites, allowing partner proteins to target complexes to specific positions. A PRE/TRE sequence requires a high density of motifs, however the composition of motifs is variable. How PRE/TRE are regulated to be activated in a tissue or time-specific manner is still under debate (Ringrose and Paro, 2007). PRE/TRE sequences have not been found in mammals so far, with the exception of two studies that described sequence-based Polycomb recruitment. A vertebrate PRE named PRE-kr was identified in mouse and showed

Table 1.3 Polycomb protein complexes

Polycomb complexes	Drosophila	Mammalian	function
PRC2	E(z)	EZH2	Histone methyltransferase containing a SET domain specific to H3K27me2/3 and H1K26me
		EZH1	Compaction chromatin
	Esc/Escl	EED1	Required for enzymatic activity of PRC2
		EED2 (PRC4)	Active in undifferentiated cells
		EED3/4 (PRC3)	Active in the absence of H1
	Su(z)12	SUZ12	Required for enzymatic activity of PRC2
	Nurf-55 (Caf1)	RbAp48/RBBP4	Histone binding proteins
		RbAp46/RBBP7	
	Pcl	PHF1	Enhancer of the H3K27 methylase activity of PRC2
		MTF2	
		PHF19	
	Rpd3	HDAC2	Histone deacetylase
	Sir2	SIRT1	H1K26 deacethylase
		JARID2	PRC2, PRC1, poised PolII recruitment
		AEBP2	Zinc finger protein. PRC2 recruitment
PRC1	dRing/Sce	RING1B/RNF2	Histone ubiquitinase H2AK119 specific
		RING1A/RNF1	
	Pc	CBX2/HPC1	Chromodomain protein reader of H3K27me3
		CBX4/HPC2	
		CBX8/HPC3	
		CBX6	
		CBX7	
	Psc	PCGF4/BMI1	Enhancer of Ring activity
		PCGF2/Mel18	
		PCGF6/MBLR	
		PCGF1/NSPc1	
	Ph	PHC1	Zinc-finger protein
		PHC2	
		PHC3	
PhoRC	Pho	YY1	Recruitment PRC2
	Phol	YY2	
	Sfmbt	L3MBTL2/ SFMBT	
RAF	Psc	PCGF1/NSPc1	Enhancer of Ring activity
		PCGF2/Mel18	
		PCGF6/MBLR	
		PCGF1/NSPc1	
	dRing/Sce	RING1B/RNF2	H2AK119 ubiquitinase
		RING1A/RNF1	
	Kdm2	KDM2B/FBXL10	H3K36me2 demethylase
		BCOR	
PR-DUB	Calypso	BAP1	Deubiquitine enzyme
	ASX	ASXL1	Enhance calypso activity

Table 1.4 Trithorax protein complexes

Trithorax complexes	Drosophila	Mammalian	function
BRM	Brm	BRM/SMARCA2	ATP-dependant chromatin remodelling. Related to the SWI/SNF complex
		BRG1/SMARCA4	
	Osa	BAF250/mARID1A/ELD	
	Mor	BAF170/SMARCC2	
	Snr1	BAF47/SMARCB1/SNF5	
ASH1	Ash1	ASH1L	H3K36 methyltransferase
	(or) Ash2	ASH2L	
	CBP		Histone acetyltransferase
TAC1	Trx	MLL1-3	H3K4 methyltransferase
	CBP		Histone acetyltransferase
	Sbf1		Antiphosphatase
MLL		MLL1/TRX1	H3K4 methyltransferase
		MLL2	
		MLL3	
		WDR5	WD repeat protein. Helps complex assembly
		ASH2L	Enhancer of MLL
		SWD1/RBBP5	
		CFP1	

affinity for PRC1 (Sing et al., 2009). An additional 1.8kb region was discovered between HOXD11 and HOXD12 and named D11.12. Interestingly, luciferase expression was reduced in a construct containing the D11.12 sequence but this required the binding site specific to YY1, a mammalian homologue of Pho, and the presence of BMI1 or EED. Ectopic integration of D11.12 led to the enrichment of BMI1, SUZ12 and H3K27me3 (Woo et al., 2010). Finally, in line with the sequence-based recruitment of Polycomb complexes, EZH2 occupies 2461 promoters in mouse ES cells and 88% of these binding sites correlate with CpG islands (Ku et al., 2008).

Although it was shown that Polycomb proteins could be recruited by specific short sequences, this mode of targeting has not been proven to be predominant in mammals. On the other hand, a variety of associated proteins could be partly responsible for Polycomb binding in vertebrates. For instance the Pho homologue YY1 associates with EZH2 in mouse undifferentiated myoblasts at silent muscle-specific genes. The use of RNAi against YY1 impaired EZH2 recruitment and H3K27me3 enrichment (Caretti et al., 2004). Two additional proteins were shown to co-localise with PRC2 at most Polycomb targets in ES cells, JARID2 and AEBP2. JARID2 contains an ARID domain that binds AT-rich sequences and is required for PRC1 and RNA Pol II recruitment (Landeira et al., 2010). AEBP2 is a zinc finger protein that binds a specific DNA motif (Kim et al., 2009).

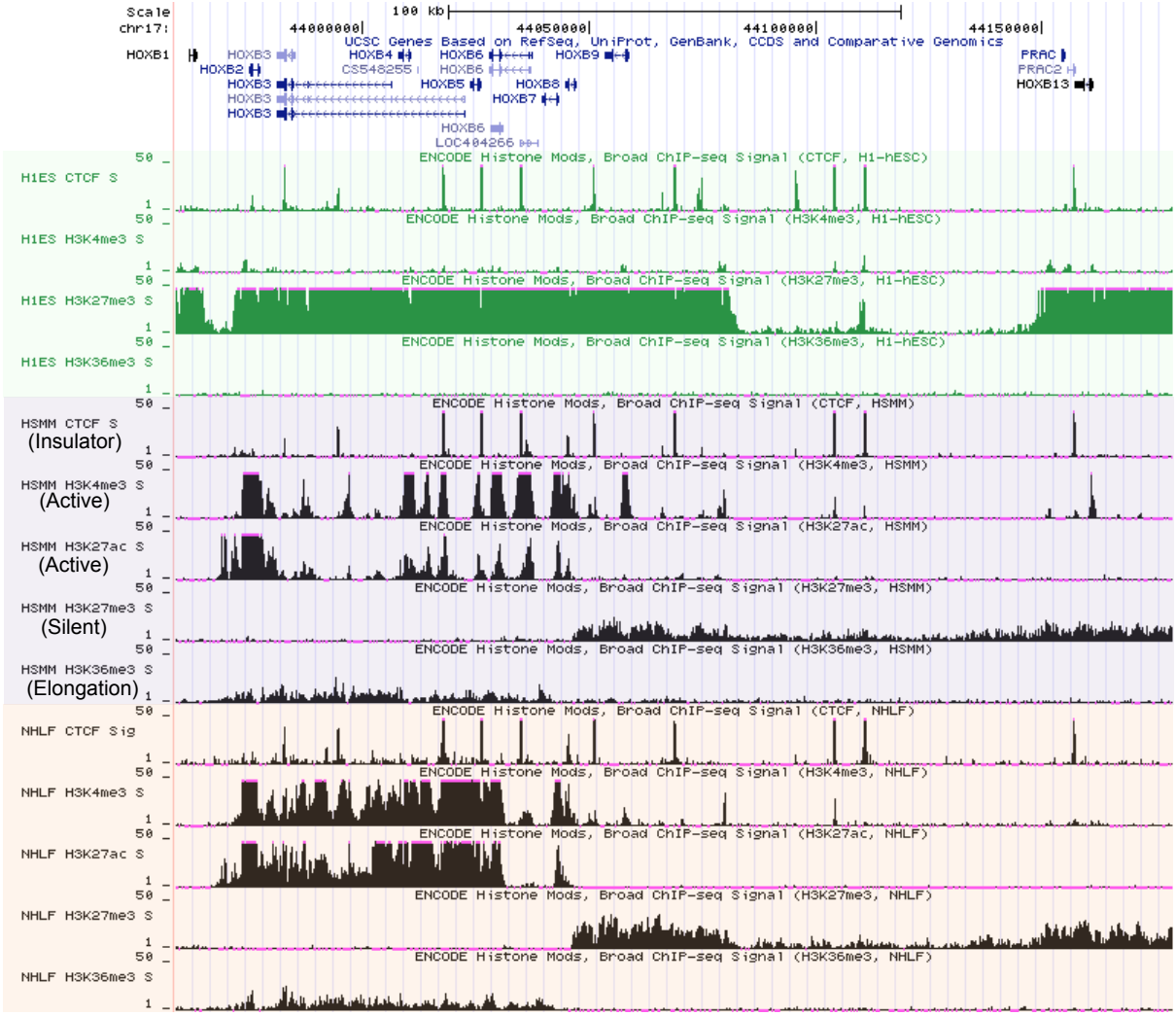
The role of ncRNA in polycomb recruitment was recently extensively studied. To date, several examples have suggested a crucial role of ncRNA mechanisms in silencing. X-chromosome inactivation in female cells is initiated by the mono-allelic

expression of a long, non-coding RNA, *Xist* that accumulates in cis on the X-chromosome. In mouse, a small 1.6kb-RNA within *Xist* named Repeat A is suspected to target PRC2 as its depletion led to a decrease in H3K27me3 (Zhao et al., 2008). Importantly, hox genes were also identified as ncRNA targets with the discovery of HOTAIR and ANRIL. HOTAIR is one of the 231 ncRNA transcribed within the four hox loci, found in human primary fibroblasts. All four hox loci were characterised for expression as well as enrichment in H3K27me3, SU(Z)12 and RNA pol II. ncRNAs were found differentially expressed in the various tissues with opposite chromatin states. HOTAIR is a 2.2kb ncRNA that was identified between hoxc11 and hoxc12, expressed in the foreskin. Abrogation of HOTAIR led to the derepression of hox genes, the decrease of H3K27me3 enrichment and the loss of Su(z)12 binding in the HoxD cluster, as HOTAIR directly interacts with SU(Z)12 (Rinn et al., 2007). It was later confirmed that the 5' domain of HOTAIR binds to PRC2 while the 3' domain binds to LSD1, responsible for H3K4 demethylation (Tsai et al., 2010). Lastly, an antisense ncRNA was identified in the Ink4b/ARF/Ink4a tumour suppressor locus and named ANRIL. The PRC1 member CBX7 was shown to bind H3K27me3 as well as ANRIL, this last interaction being essential for gene repression (Yap et al., 2010).

Hox genes were identified as bivalent domains (Bernstein et al., 2006) and large-scale studies have confirmed the bivalent state of hox gene clusters (**Figure 1.11**) (Raney et al., 2011). Polycomb and trithorax proteins play a major role in the regulation of those genes. However, their mechanisms of action are not fully understood.

It is generally accepted that PRC2 generates high levels of H3K27me3 and provides a docking site for PRC1 which then catalyses mono-ubiquitination of H2A. Yet, it should be noted that two types of bivalent domains were identified, some carrying both PRC2 and PRC1 and some carrying PRC2 only, suggesting different mechanisms of action. It was observed that PRC1-positive domains are better conserved between mouse and human, correspond to developmental regulators and retain more H3K27me3 upon differentiation (Ku et al., 2008). As previously mentioned, the PRC1 member RING1B generates mono-ubiquitination of H2A that recruits RNA pol II in a poised state and therefore blocks transcription (Stock et al., 2007). Furthermore, the cooperative complex dRAF brings to action the component KDM2B that specifically demethylates H3K36me2, a mark associated with transcription elongation (Lagarou et al., 2008). In addition, when PRC1 was characterised, it was shown to compete and block the action of the remodelling complex SWI/SNF (Shao et al., 1999). In fact, electron microscopy later confirmed that a recombinant PRC1 complex was able to compact nucleosomal arrays (Francis et al., 2004). Another member of PRC1, BMI1, was identified as an important player in the silencing process by interacting with DNMT associated protein (Dmap) and DNA methyltransferase1 (Dnmt1), responsible for CpG methylation generally associated with repressed transcription. Knockdowns of BMI1 or Dmap in MEFs led

Figure 1.11 Hox genes ENCODE data



The ENCODE (Encyclopaedia of DNA elements) project maps functional elements all over the human genome. Here are represented ChIP-seq data obtained for the Hoxb cluster including CTCF occupancy along with active marks (H3K4me3; H3K27ac), the elongation mark H3K36me3 and the silencing polycomb mark (H3K27me3). Three cell lines were compared: human ES cells (green), human skeletal muscle myoblasts (purple) and human lung fibroblasts (orange).

In ES cells, the hoxb genes are silent and the hoxb cluster is mostly occupied by large blankets of H3K27me3. Peaks of H3K4me3 were detectable in some promoters but the elongation mark H3K36me3 was totally absent. Strikingly in differentiated cells, H3K4me3 and H3K27me3 became almost completely exclusive marks. Along with H3K4me3, the enrichment of H3K27ac and H3K36me3 increased, suggesting active transcription at those loci.

to hox loci derepression (Negishi et al., 2007). Similarly, the PRC2 member EZH2 co-precipitates with Dnmt1, 3a and 3b, and has been shown to target Dnmts and induce CpG methylation at promoters (Vire et al., 2006). Finally, the H3K4 demethylase RBP2 was co-purified with PRC2 members and shown to co-localise with SUZ12 and EZH2 at PRC2 target genes. PRC2 recruits RBP2, as knockdown of SUZ12 decreased EZH2 and RBP2 recruitment. On the other hand, induction of EED increased EZH2, SUZ12 and RBP2 recruitment as well as H3K27me3 enrichment while H3K4me3 levels were reduced (Pasini et al., 2008).

An elegant model system using ES cells, neuronal progenitors and glutamatergic pyramidal neurons allowed the analysis of DNA methylation, polycomb and trithorax mark enrichments as well as RNA pol II binding and activity during differentiation. As expected, DNA methylation levels increased on the promoters of genes associated with pluripotency or for instance the germline. However a small proportion of promoters showed a decrease in DNA methylation, correlating with RNA pol II binding and loss of H3K27me3 enrichment. This small group of genes was usually involved in neuronal development, ion transport or neurotransmitter regulation (Mohn et al., 2008). In another study, the HoxD cluster was analysed at different stages of mouse-tail bud development. As late hox genes were induced, RNA pol II binding, H3 acetylation and H3K4me3 increased while enrichment of H3K27me3 was lost (Soshnikova and Duboule, 2009). It was therefore confirmed that in order to get lineage-specific gene expression during differentiation, bivalency has to resolve into one or the other mark.

The mechanisms described so far for polycomb silencing involved the inhibition of chromatin remodelling and transcription elongation and recruitment of Dnmts. In

contrast, trithorax group proteins act during differentiation to allow the transcription of target genes. In fact, MLL1 is required for neurogenesis in mouse postnatal brain. If MLL1 is knocked out, the neurogenesis regulator Dlx2 is not induced and is occupied by H3K27me3 in the subventricular zone contrary to what was observed in wild type mice (Lim et al., 2009). Under retinoic acid treatment, the human demethylase UTX interacts with the MLL complex and specifically targets H3K27me2/3. Indeed, RA treatment leads to increased UTX occupancy and H3K4me3 enrichment as well as a diminution of H3K27me3 and SUZ12 binding. siRNA against UTX raised H3K27me2/3 levels and PRC1 occupancy on several hoxa and hoxc genes, correlating with loss of expression (Lee et al., 2007b). Similarly, JMJD3, a H3K27 demethylase was shown to associate with the MLL complex. During macrophage transdifferentiation, NFkB activates JMJD3 that then demethylates polycomb target genes including late Hoxa (De Santa et al., 2007). Like Polycomb complexes, Trithorax proteins influence DNA methylation levels. In fact, the CXXC domain of MLL binds to unmethylated CpG and prevents DNA methylation as observed on the hoxa9 locus (Erfurth et al., 2008).

Finally, the active transcription-associated histone variant H2A.Z seems to play a role in bivalency. H2A.Z colocalises with SUZ12 in undifferentiated cells but during differentiation, induced genes become enriched in H2A.Z and lose the polycomb mark (Creyghton et al., 2008). H2A.Z was also shown to inhibit DNA methylation in *Arabidopsis thaliana* (Zilberman et al., 2008).

1.4.3.3 Maintenance of Polycomb/Trithorax marks

One of the big challenges of epigenetic research is to understand how epigenetic patterns are transmitted from one cell to another during mitosis. The passage of the replication machinery and extreme chromatin condensation during mitosis are both likely to disrupt DNA/protein interactions and histone mark distribution. However, polycomb and trithorax complexes are indispensible for the long-term maintenance of hox gene expression and therefore must be retained on allocated target loci after cell division. Little is known about the inheritance of Polycomb/Trithorax marks but several studies have described possible mechanisms.

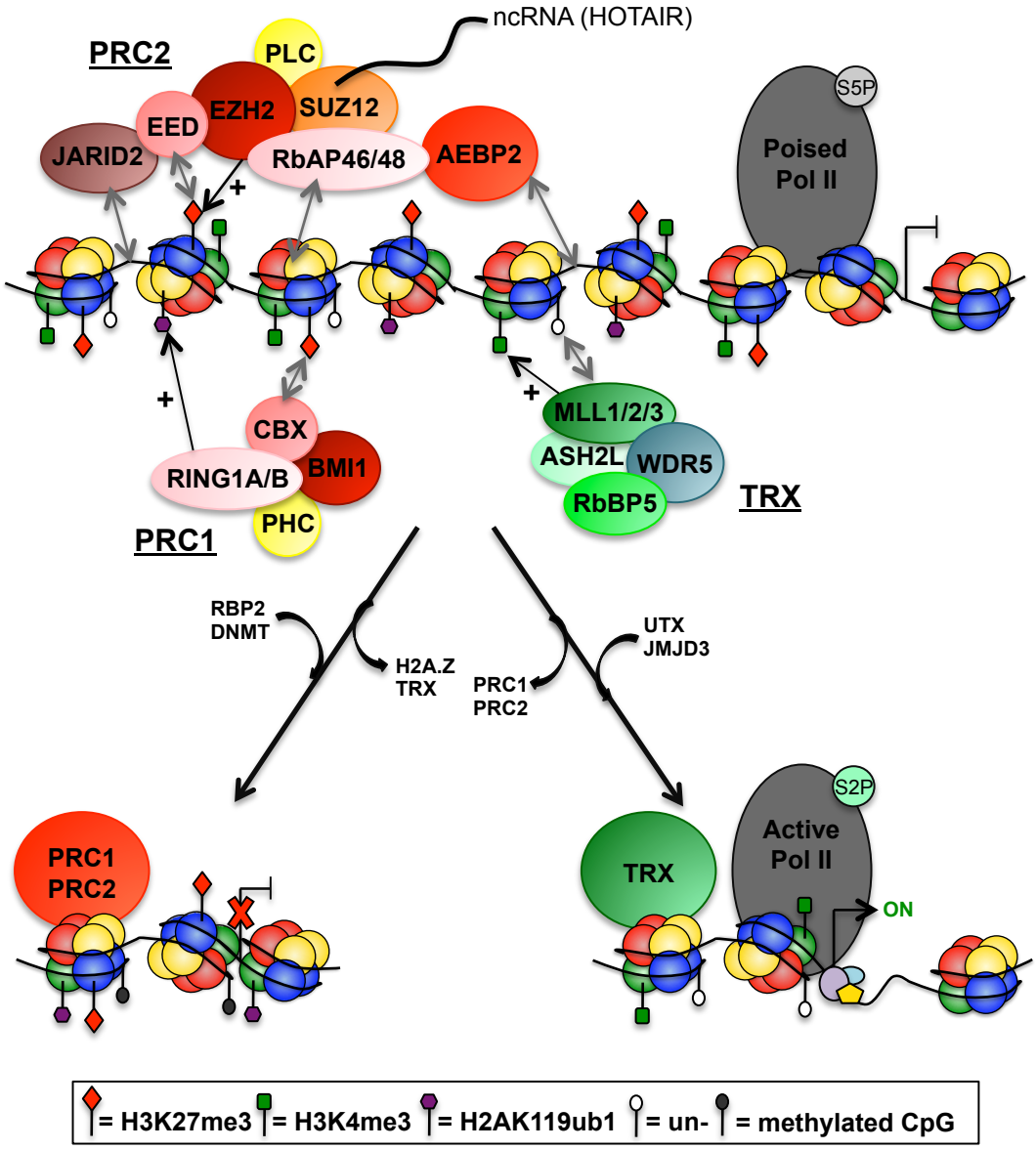
In the case of a DNA sequence-based recruitment, it would be easy to imagine the role of specific motifs in Polycomb/Trithorax retargeting after mitosis. PRE/TREs have been identified in *Drosophila* and are believed to play a major role in memory. But PRE/TREs have been only rarely found in mammals so are unlikely to provide a memory system. Instead, targeting factors that associate with complexes or ncRNA could help binding to chromatin after cell division. In addition, CpG islands are suspected to play a role in inheritance, as this is where most PRC2 members bind. However, it is important to note that only a minority of CpG islands bind EZH2 and H3K27me3, suggesting the need for other factors to retain Polycomb binding (Ku et al., 2008).

PCNA (proliferative cell nuclear antigen) is an important protein that helps progression of the DNA polymerase during replication. Immuno-fluorescence staining revealed colocalisation of EZH2, PCNA and BrdU in human fibroblasts during S phase. Moreover, the study showed that EZH2, EED and SUZ12 together bind H3K27me3, providing a new model where the polycomb mark is copied on new

histone during DNA replication. When a construct containing five Gal4 binding sites linked to a luciferase reporter was stimulated by GAL4-EED, enrichment of EED, EZH2, SUZ12 and H3K27me3 were raised while RNA pol II binding was lost. Importantly, six days following GAL4-EED washout, luciferase remained silent and levels of H3K27me3 stayed high. This phenomenon was abolished when siRNAs were used against EZH2 or SUZ12 (Hansen et al., 2008). The PRC1 core complex was shown to bind a DNA template containing a replication origin and be still associated to it after replication. Moreover, when competitor DNA was incubated with PRC1-prebound templates, PRC1 did not relocate during replication. Finally, ChIP was performed every 10 minutes during S phase and levels of Psc and Pc were constant at three different PREs of the Bithorax complex (Francis et al., 2009). Constant binding to chromatin of PRC2 members was confirmed in mouse embryonic fibroblasts. Cytological approaches revealed constant H3K27me3 enrichment during all phases of mitosis as well as SUZ12 and EZH2 colocalisation with the polycomb marks. However, staining of PRC1 member BMI1 became very weak during mitosis and progressively re-localised with H3K27me3 during G1 (Aoto et al., 2008). Dissipation of BMI1 during S-phase had already been observed and is associated with BMI1 phosphorylation (Voncken et al., 1999).

Unlike SETD1A, Pol II or LSD1, the two subunits MLL-C and MLL-N colocalise with condensed chromatin and promote rapid reactivation of target genes after mitosis. ChIP experiments also highlight the presence of other members of the MLL complex during M-phase, including RbBP5 and ASH2L. Finally, when MLL was knocked down, target genes were less rapidly activated after mitosis (Blobel et al., 2009) (**Figure 1.12**).

Figure 1.12 Regulation of "Bivalency"



Polycomb and Trithorax proteins regulate many genes in ES cells and during development, including the Hox genes. Polycomb members EZH2 and RING1A/B deposit the silencing-associated marks H3K27me3 and H2AK119ub1 whereas Trithorax member MLL deposits the activating mark H3K4me3. Multiple associated proteins and non-coding RNA (ncRNA) help target the complexes to specific genes. In ES cells, poised genes are enriched for both active and repressive modifications a state named "bivalency". During development, bivalency often resolves into one or the other mark according to the loci's transcriptional status. (Single-headed arrows indicate mark deposition and double-headed arrows indicate interaction).

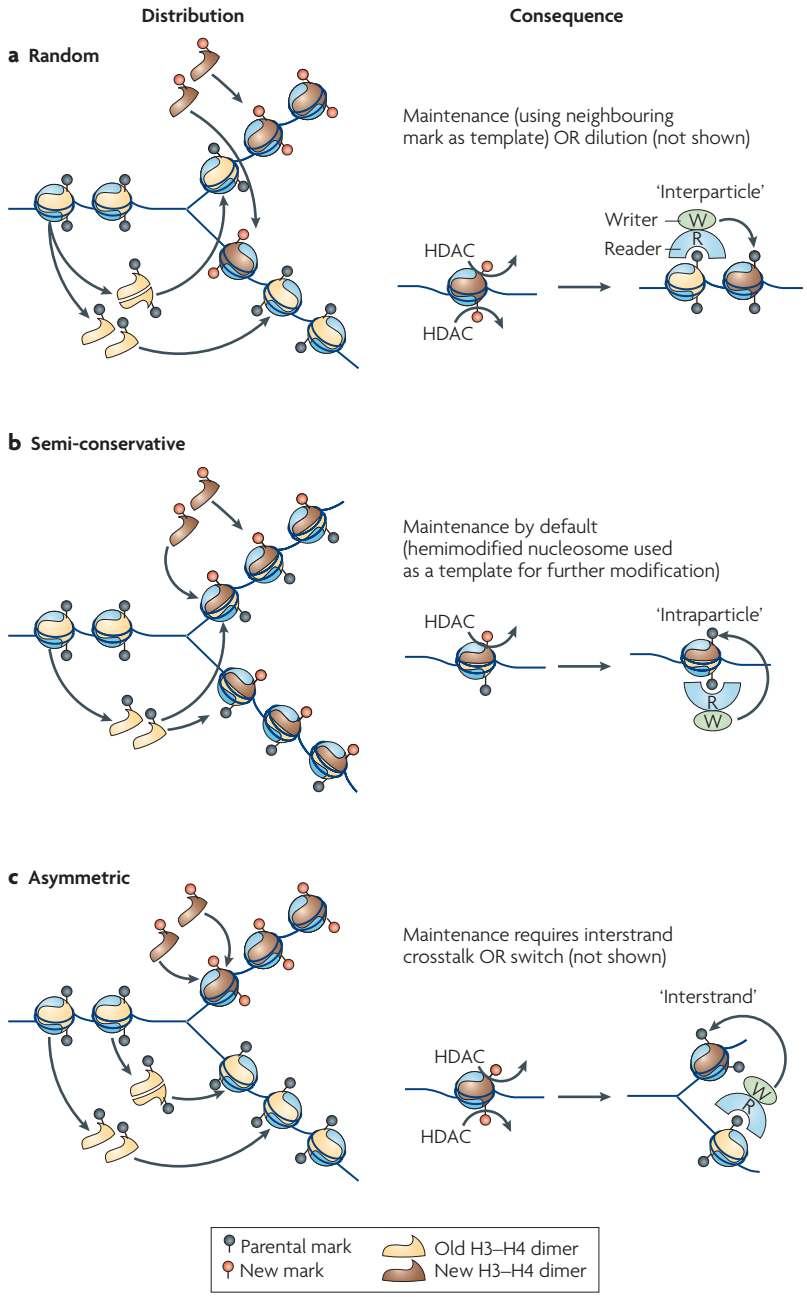
1.5 Epigenetic inheritance

A widely accepted definition of epigenetic phenomena says that it corresponds to “a change in the state of expression of a gene that does not involve a mutation but that is nevertheless inherited in the absence of the signal that initiated the change”. Therefore, epigenetic modifications like histone marks or DNA methylation should survive DNA replication and cell division. However, little is known about how chromatin states are transmitted between cell generations, and whether histone modifications are actively replicated is still under debate (Ptashne, 2007).

1.5.1 Chromatin assembly

Replication is the process by which the double-stranded DNA molecule is faithfully duplicated in order to produce enough genetic material for both daughter cells after mitosis. We now know that the DNA molecule assembles into chromatin through its association with histone proteins. Therefore, not only the genomic sequence has to be copied during replication but the all of the chromatin fibre in its specific state. Replication leads to the dilution of preassembled histones via the incorporation of newly synthesised histones. DNA replication progresses in a semi-conservative manner where each new double strand DNA contains one mother strand. However, how old and new histones are distributed during replication is still a mater of debate. Three putative models have been described: (1) “old” histones are randomly distributed between the two new chromatin fibres, (2) “old” histones are distributed in a semi-conservative manner so the two halves of a nucleosome go to different nascent fibre and (3) the “asymmetric model” in which all H3/H4 tetramers go to one chromatin fibre or the other (**Figure 1.13**) (Probst et al., 2009).

Figure 1.13 Distribution of "old" and "new" H3/H4 dimers during replication



The fate of "old" and "new" H3/H4 histone dimers during replication is not well understood. Three potential mechanisms have been described so far. Parental histones could be randomly distributed on both daughter strands (**a**) or equally distributed according to the semi-conservative model (**b**). Finally, the asymmetric model predicts the old histones to incorporate exclusively on one daughter strand and therefore requires interstrand crosstalk for maintenance of modifications (**c**).
In fact, in each model, parental modifications are recognised by reader proteins that then recruit chromatin modifiers for the deposition of appropriate marks on newly incorporated histones.
Taken from Probst et al., 2009

Chromatin assembly during replication involves a wide variety of histone chaperones that help displace old histones and carry newly synthesised ones. How they proceed exactly remains unclear but it is now accepted that they play an important role in chromatin assembly during replication, transcription or DNA repair (Avvakumov et al., 2011). An elegant study using HeLa cells expressing flag-tagged H3.1 allowed the classification of events necessary to transport newly synthesised histones into the nucleus. H3.1 is a H3 variant exclusively transcribed during S-phase and is incorporated during replication. Fractionation and mass-spectrometry found four complexes which were classified in their probable order of action: (I) contains the histone chaperone HSP70 and H3.1 that carries K9me1, (II) contains HSP90, another chaperone tNASP, H4 and H3.1K9me1, (III) is the most abundant complex and includes the chaperone sNASP, H3.1, H4K5acK12ac, the histone acetyl transferase HAT1 and RbAp46 and finally (IV) is thought to be involved in nuclear import as it contains H3.1, H4, Importin-4 and the histone chaperone Asf1 (Campos et al., 2010). The role of modifications carried by new histones has not been elucidated. One possibility is that they increase the affinity of specific histone chaperones (Avvakumov et al., 2011).

In the context of DNA repair, Asf1 was shown to transiently associate with one of the three subunits of chaperone Caf-1, p60 before nucleosome assembly (Mello et al., 2002). Studies on replication suggest the chaperone CAF-1 plays a central role in histone deposition. During G1, the pre-replication complex binds to DNA at replication origins. Additional factors and activated CDKs then follow and associate with the helicase complex of MCM proteins and PCNA, a ring-like trimer that encircles the DNA and drags the DNA polymerases during replication (Moldovan et

al., 2007). Importantly, the Caf-1 subunit p150 was shown to interact with PCNA and both proteins colocalise at replication loci. Therefore, Caf-1 was strongly suspected to play a crucial role in newly synthesised histone incorporation during replication (Shibahara and Stillman, 1999).

Asf1 is involved in the transport of new histones but also in the transport of parental histones as it was shown to interact with histones carrying modifications uncharacteristic of newly synthesised histones. In addition, this chaperone interacts with the helicases MCM2, 4, 6 and 7 through an H3/H4 bridge. When Asf1 was knocked out, a block in S-phase was observed, along with helicase activity defects. These results strongly suggest that Asf1 acts as an acceptor and donor of parental and new histones during replication through its interaction with the MCM complex (Groth et al., 2007). Finally, specific histone chaperones including Nap1, FACT or Chz1 are involved in the transport of H2A-H2B dimmers (Ransom et al., 2010).

Because H2A/H2B are located at the periphery of the nucleosome, H3 and H4 assemble first and disassemble last. However, the composition of newly formed H3/H4 tetramers is still unknown (Ransom et al., 2010). sNASP was shown to dimerise and form a complex with two H3/H4 dimers. Therefore, one could think that Asf1 imports new H3/H4 tetramers into the nucleus (Campos et al., 2010). Yet, analysis of Asf1-H3-H4 stoichiometry and X-ray crystallography revealed that the complex exists as a stable heterotrimer and even blocks the formation of a H3/H4 tetramer (English et al., 2006; English et al., 2005). It was therefore suggested that Asf1 is involved in parental H3/H4 tetramer splitting and dimer distribution, which provides an argument for the mixing of new and old histones into newly formed nucleosomes. The semi-conservative model of segregation is therefore an attractive

one. Modifications on the parental histones would be bound by modifiers that could then deposit the appropriate mark on the new histones. However, this model does not fit with the fact that the daughter fibres are not synthesised at the same speed, and gives less opportunity for change. On the other hand, the random distribution of nucleosomes does not explain how a specific mark can stay restricted on one or two nucleosomes (Probst et al., 2009).

1.5.2 DNA methylation

DNA methylation is a well characterised modification associated with gene silencing. It plays a major role in gene imprinting, X-inactivation, silencing of retrotransposons or carcinogenesis (Watanabe and Maekawa, 2010). Importantly, patterns of DNA methylation and histone modifications have been positively or negatively correlated. For instance, CpG-rich promoters are generally enriched in H3K4me3 and depleted in DNA methylation (Mikkelsen et al., 2007). DNA methylation was shown to coincide with long regions of H3K9me2, particularly in pericentromeric chromatin, in *Arabidopsis thaliana* (Bernatavichute et al., 2008). Because DNA methylation is a more stable modification, a model was proposed in which DNA methylation provides a platform for histone modification inheritance (Blomen and Boonstra, 2011).

BMI1 (PRC1) and EZH2 (PRC2) were shown to interact with Dnmt1 (Negishi et al., 2007; Vire et al., 2006). It is widely accepted that Dnmt1 is the enzyme that maintains DNA methylation levels during replication. However, a study using ES cells knockout for dnmt1, 3a or 3b revealed a role for Dnmt3a/3b for the maintenance of methylation states in CpG poor regions. In fact, in the absence of Dnmt1, specific sequences,

including the promoters of LINE-1 repetitive elements remained methylated, but this was lost when Dnmt3a/3b were absent. This suggest that de novo methylation by Dnmt3a/3b could compensate for the poor maintenance by Dnmt1 (Liang et al., 2002). Nevertheless, Dnmt1 was shown to rapidly methylate CG specifically on hemimethylated DNA (Hermann et al., 2004). Restoration of methylation levels after 5-azacytidine treatment was impaired in the absence of Dnmt3a/3b which confirmed that Dnmt1 preferentially acts on hemimethylated sequences (Liang et al., 2002).

Importantly, Dnmt1 binds to PCNA at replication foci (Chuang et al., 1997) and the SET and Ring finger-associated (SRA) domain protein Np95 was shown to accumulate at DAPI-dense heterochromatin regions in mid-to-late S-phase in ES cells. Immunoprecipitation of Np95 revealed its association with Dnmt1 and PCNA. Np95 and Dnmt1 interact, such that in Dnmt1/3a/3b triple knockout cells, where DNA methylation is totally absent, Np95 did not localise at heterochromatin. However in Dnmt1 knockout cells, Np95 could localise to heterochromatin but importantly, binding was independent of the cell cycle phase, suggesting that Np95 prefers hemi-methylated DNA. In Np95 knockout cells, chromatin was generally hypomethylated and genes derepressed. In addition, Dnmt1 showed a diffuse distribution throughout the cell cycle, suggesting that targeting of Dnmt1 required interaction with Np95 (Sharif et al., 2007).

1.5.3 Heterochromatin maintenance

The heterochromatin protein HP1 was shown to bind H3K9me2/3, a histone mark catalysed by the methyl-transferases SUV39H1 or SETDB1 (Bannister et al., 2001). During mitosis, a transient increase in H3S10p and H3K14ac levels induce the

eviction of HP1 (Fischle et al., 2005). However, the CAF-1 subunit p150 was shown to interact with HP1, suggesting a role of this histone chaperone in heterochromatin maintenance. Consistent with this, abolishment of the p150-HP1 interaction led to a cell cycle arrest during pericentric heterochromatin replication (Quivy et al., 2008). Furthermore, p150/CAF-1 interacts and colocalises with the methyl-CpG binding protein MBD1 in a complex containing HP1, providing a link between heterochromatin maintenance and DNA methylation (Reese et al., 2003). Indeed, it should be mentioned again here that p150/CAF-1 interacts with PCNA (Shibahara and Stillman, 1999). Finally, PCNA was shown to interact with many proteins that play a role in heterochromatin formation, including HDACs. HDACs themselves bind to Dnmt1, reinforcing the role of DNA methylation (Probst et al., 2009).

Although DNA methylation seems to be important for heterochromatin formation, several studies highlight the fact that histone modifying enzymes are crucial for the maintenance of DNA methylation. In mouse embryonic fibroblasts, when the histone methyltransferase Suv39h was knocked out, Dnmt3b interaction with HP1 at heterochromatin foci was abolished. In addition, DNA methylation was specifically altered at pericentric major satellite repeats. However, H3K9me3 was not impaired in dnmt1 or dnmt3a/b knockout ES cells (Lehnertz et al., 2003). Furthermore, G9a knockout in ES cells induced a loss of DNA methylation at retrotransposons, major satellite repeats and CpG rich promoters. However, no decrease in H3K9me3 or HP1 binding was observed and these regions remained silent. In addition, DNA methylation was rescued when a catalytically inactive G9a was transfected in knockout ES cells, suggesting that H3K9me3 itself is not responsible for DNA methylation at these loci (Dong et al., 2008).

1.5.4 Transcriptional memory

Several studies have implicated nuclear organisation and specific histone variants in the propagation of active transcription. For instance, the Ino1 gene is activated under inositol starvation and the Gal1 gene is activated under galactose exposure of yeast. During activation, these genes are relocated to the nuclear periphery and interact with the nuclear core component Nup2. After the genes are repressed, they stay at the nuclear periphery for several cell cycle rounds and show much faster reactivation when stimulated. Importantly, this phenomenon of transcriptional memory required intact H2A.Z (Brickner et al., 2007). Gene loops that juxtapose a promoter and a terminator of transcribed genes were identified in yeast and mammals. Using the 3C technique, the authors monitored gene looping in different Gal genes during induction, repression and reinduction. When Gal10 is repressed, RNA pol II dissociates, whereas looping persists for at least 4 hours. Interestingly, in the Sua7-1Δ strain where looping is abrogated, no effect was observed during initial activation but the kinetics of reactivation of Gal10 or Gal1p-Sen1 was impaired. This observation correlated with a less rapid reassociation of RNA pol II during reinduction and loss of Gal4. A role for looping in nuclear pore anchoring was therefore suggested (Laine et al., 2009). Another study confirmed that transcription induction by galactose exposure of a naïve Gal1 gene was slower than induction of a recently transcribed Gal1 (Kundu et al., 2007). This rapid reinduction was observed even after cell division. ChIP experiments showed that, although all components of the transcriptional machinery disappeared during glucose repression, RNA polII and TBP were recruited much faster during reinduction compared with initial induction. Analysis of reinduction kinetics in yeast mutants revealed that histone modifications

like H3K4me, H3K36me, H3K79me, H3ac or H2Bub are not essential for transcriptional memory. However, the SWI/SNF complex member snf2 was shown indispensable (Kundu et al., 2007).

Although PREs have been rarely identified in mammals, their role in transcription memory should be noted. The scalloped (sd) gene, important for wing development in *Drosophila*, was repressed by the upstream insertion of the PRE element Fab-7. This silencing was the result of an interaction between the inserted and endogenous Fab-7, as when the endogenous Fab-7 was mutated, sd was derepressed. Moreover, this phenomenon was cancelled by polycomb mutations, suggesting that chromosome pairing stabilised polycomb silencing. Importantly, when flies were back-crossed in order to obtain repressive phenotypes, the derepression and loss of pairing remained for at least five generations. This result showed that nuclear architecture plays an important role in the inheritance of polycomb silencing (Bantignies et al., 2003). Another study using FISH and 3C confirmed that silencing in the BX-C locus requires many interactions between PRE and promoters in *Drosophila*. This phenomenon was observed in embryos and in ES cells. Interestingly, the PRE Fab-7 was shown to lose its interaction with other element in a cell line where AbdB was transcribed. When the PRC1 protein Pc was knocked down, interactions between the three bithorax genes was impaired. A month following restoration of Pc levels, interactions were almost completely recovered. Interestingly, RNAase treatment of cells did not affect interactions, suggesting no role for RNAs (Lanzuolo et al., 2007).

H3.3 had been shown previously to mark zones of active transcription. Therefore, a model of transcriptional memory was suggested where H3.3 is incorporated during active transcription and then inherited during replication (Ahmad and Henikoff, 2002). Interestingly, H3.1 and H3.3 carry different post-translational modifications. In their cytosolic and nuclear forms, they were mostly unmodified, however, H3K9me1 enrichment was detected on H3.1 whereas H3K9ac was found on H3.3. This is believed to potentiate the action of the methyl-transferase SUV39H1 on H3.1 as H3K9me1 is a better substrate. As expected, nucleosomal H3.3 was associated with active marks in contrast with H3.1. Furthermore H4 associated with H3.3 were more acetylated than that associated with H3.1. Finally, in an array containing neighbouring H3.3 and H3.1, active marks were detected on H3.1 variants, suggesting that H3.3 enhances active chromatin states (Loyola et al., 2006). The purification of tagged H3.1 and H3.3 variants and the use of mass spectrometry revealed two different pathways for the incorporation of these histones. H4, ASF1, NASP, HAT1, p48/CAF-1 and importin-4 were identified as common components of H3.1 and H3.3 complexes. On the other hand, p150/CAF-1 and p60/CAF-1 were only found to be associated with H3.1 whereas the HIRA chaperone was exclusively found in the H3.3 complex. Using UV-induced DNA damage, the authors showed that, in contrast with H3.1, the H3.3 variant is incorporated in a DNA-synthesis-independent manner. In addition, both variants never associated in the same nucleosome (Tagami et al., 2004).

1.6 Aims

“Epigenetic” modifications are commonly defined as events that control gene expression by mechanisms other than change in the DNA sequence. However, this term also includes the expectation that these marks are inherited as they are strongly suspected to define cell identities (Holliday, 2006). If histone modifications are truly epigenetic marks, they should therefore (1) be causative for transcription and (2) be inherited throughout the cell cycle, although the signal(s) that induced them has disappeared. Histone modifications have been widely correlated with transcription. However, it has not been clear so far if they are responsible for, or only a consequence of, transcriptional status. Furthermore, histone modifications are suspected to play a role in cell memory by providing a docking site for modifying enzymes during replication. Because modifying enzymes often bind to the mark they deposit, a simple model proposes that they recognise marks on parental histones and deposit them on the adjacent newly assembled nucleosomes.

Therefore in this study, we decided to test the epigenetic nature of histone modifications and more specifically histone acetylation. Using an HDAC inhibitor, we design different model systems in order to:

- Measure the impact of hyperacetylation of histone on chromatin states and transcriptional activity.
- Study the long-term effects of a transient increase in histone acetylation.

2. Materials and Methods

2.1 Mouse Embryonic Stem (ES) Cell Culture

2.1.1 CCE/R cells

CCE/R cells are an adherent male ES cell line. Cells were grown on gelatine-coated flasks (T25/75 cm^2, Sarstedt) at 37°C in humidified air with 5% CO_2 in Dulbecco's modified eagles medium (DMEM, Gibco). The culture medium was supplemented with 20% foetal calf serum (Gibco), 1% penicillin/streptomycin (Gibco), 1% L-glutamine (Gibco), 1X non-essential amino acids (Gibco), 0.25% 2-mercaptoethanol (Gibco) and 1U/μl Leukaemia inhibitory factor (LIF/ESGRO, Chemicon) (supplemented ES cells DMEM).

Cells were split every two days, usually in a 1 to 6 ratio but this varied according to cell density. Briefly, cells were dislodged from the flasks using trypsin-EDTA 1X (Gibco). Trypsinised cells were then pipetted into the old medium in order to deactivate the enzyme, centrifuged for 5 minutes at 250g and resuspended in warm fresh medium. New flasks were coated with 0.1% gelatine for at least 20 minutes before the excess was removed and the cells added.

On alternate days, cells were supplemented with warm fresh medium.

2.1.2 Differentiation of ES cells

Cells were induced to differentiate by following a method adapted from (Chambeyron and Bickmore, 2004). Cells were trypsinised and replated in a non-adherent plastic dish (Sterilin), in supplemented ES cells DMEM without LIF.

In order to induce the expression of Homeobox (Hox) genes, 1μM Retinoic Acid (RA) was added into the medium after 2 days of differentiation, and until the end of the

experiment. After 8 days of differentiation cells were replated onto adherent plastic dishes (Corning) allowing the embryonic bodies to settle and continue differentiation.

2.1.3 Eed/Ring1B double knock-out (dKO) ES cells:

Cells were generated and provided by Martin Leeb, a PhD student in Anton Wutz's laboratory (The Wellcome Trust Centre for Stem Cell Research, University of Cambridge). The dKO ES cell line was derived by conditional deletion of Ring1B in Eed-/- ES cells using Tamoxifen inductible CreER transgene (Leeb et al., 2010). Both parental wild type ES cell and dKO ES cell lines were grown on a mono-layer of irradiated Mouse Embryonic Fibroblasts (MEFs) as feeder cells.

2.1.3.1 Isolation of MEFs

MEFs were extracted from E12 mouse embryos (provided by Graham Anderson, University of Birmingham) of which the head and internal organs had been removed. The bodies were cut into small pieces in a dish with a sterile scalpel. Trypsin was added and using a syringe and needle, cells were separated as much as possible. Fresh medium was then added in order to deactivate the enzyme and the mixture transferred into a 50ml centrifuge tube where the cells were allowed to settle. The supernatant was removed and replaced by fresh media. Cells were finally transferred in a pre-gelatinised flask. Primary MEFs were cultured in DMEM supplemented with 10% foetal calf serum, 1% penicillin/streptomycin, 1% L-glutamine, 1X non-essential amino acids and 0.25% 2-mercaptoethanol (all from Gibco).

Cells were split when 100% confluence was reached and fed according to the colour of the media.

2.1.3.2 Irradiation of MEFs

Cells were trypsinised and centrifuged for five minutes at 250g. Cells were then resuspended in fresh medium to 10^7 cells/ml and irradiated for 15 mins. The radionuclide Caesium-137 was used as a source of radiation at 2.464 grays/min. Following irradiation, 10^6 and $3x10^6$ cells were respectively replated in a T25 and T75 pre-gelatinised flask. Because of their fragile nature, dKO cells were cultured on a double density MEF layer.

2.1.3.3 Culture of ES cells on feeders

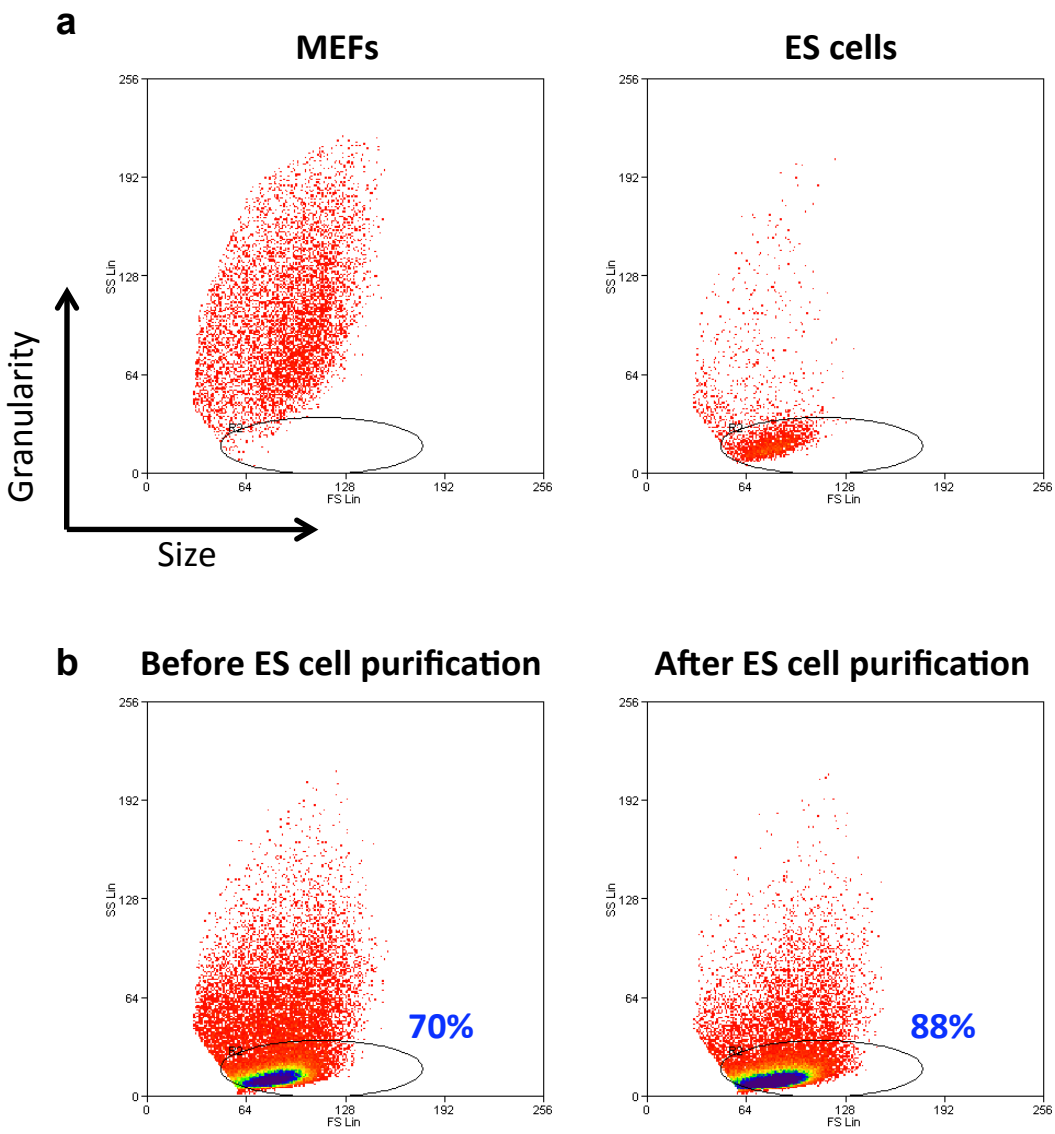
Cells were cultured on feeder cells as they were on pre-gelatinised flask with the same incubation conditions and medium. Again, cells were split every two days depending on cell confluence and fed on alternate days. When split, ES cells and MEFs were trypsinised and replated on a new layer of feeders. Most of the old MEFs wouldn't attach to the new ones and were washed out the next day with fresh medium.

2.1.3.4 Harvesting of ES cells grown on feeders

Because ES cells were attached to the MEFs when grown, it was necessary to purify them as much as possible prior any experiment. In theory, MEFs attach to the flask much faster than ES cells. Therefore, cells were trypsinised and the cell suspension replated on a pre-gelatinised flask in order to selectively remove the MEFs.

To test the purification efficiency FACS analysis was used. The yield of ES cell purification would depend on how long the cells were replated for. In order to get the purest ES cell population with limited ES cell loss, we decided to replate the cells after trypsinisation for different periods of time and analyse the proportion of each population in the final cell suspension. Initially, MEFs alone and ES cells grown on MEFs were analysed so as to gate and quantify each population on the two-dimensional scatter plot. While MEFs revealed a wide range of granularity and size, ES cells appeared to be morphologically more homogeneous. However for our experiment, it was possible to distinguish the two types of cells on the plots **(Figure 2.1.a)**. ES cells that had been grown on MEFs were then harvested and replated for 20, 30, or 45 minutes. Some protocols advised to replate cells for shorter times but consecutively in different pre-gelatinised flasks. Therefore we also analysed cells after replatment twice for 15 minutes or 3 times for 15 minutes. The percentage of MEFs and ES cells in the final cell suspension for each condition were given. The maximum purity we were able to reach was 88% ES cells after 20 minutes in a single pre-gelatinised flask **(Figure 2.1.b)**. When purification was carried out for longer, loss of ES cells progressively increased. However, when cells were replated for shorter times in multiple flasks the purity rose to approximately 87% **(Table 2.1)**. We thus decided for our experiments to replate ES cells that had been grown on feeders for 20 minutes after trypsinisation.

Figure 2.1 Light scatter plots of ES cells and MEFs



a, Two-dimensional light scatter plots were obtained passing the cells through a flow cytometer. On the X-axis, forward light scatter indicates the size of the cells while the side light scatter on the Y-axis indicates the granularity of the cells. First, irradiated MEFs alone were analysed and the corresponding population could be seen on the plot. Subsequently, ES cells grown on MEFs were analysed and as expected a new population appeared on the plot. A gate was used to target the ES cell population and permit quantification.
b, Because ES cells were grown on a layer of MEFs, it was essential to purify them as much as possible prior any experiments. For that, cells were trypsinised and replated on a pre-gelatinised flask for 20mins. Because MEFs attach to the flask before ES cells, the media was progressively depleted in MEFs. Cell populations were analysed prior and after purification. % indicate proportion of stem cells

Table 2.1 ES cell purification Yield according to the replatement time

Time of replatment	Final % of ES cells	Final % of MEFs
0 min	74	26
20 mins (no gelatine)	78	22
20 mins	88	12
30 mins	84	16
45 mins	83	17
2 x 15 mins	88	12
3 x 15 mins	87	13

When ES cells grown on MEF feeders were harvested by trypsinisation, the resultant cell suspension contained both cell lines. It was therefore important to separate the two cell types as much as possible before any experiment. In order to optimise the harvest of ES cells grown on feeders, FACS was used to calculate the yield of ES cell purification after replatement of the cell suspension on gelatine for different periods if time. Along with purification, we observed a significant loss of cells. In fact in a 20 minutes time, a substancial number of ES cells had attached to the flask. Thus, replatement in a flask without gelatine was tested.

2.2 Cell culture monitoring

2.2.1 Cell death and Proliferation assay

2.2.1.1 Cell count

Cell viability was determined using a Malassez hemocytometer. Cells were stained with Trypan blue and those that excluded the dye were considered viable whereas dead cells stained blue.

2.2.1.2 Measure of proliferation

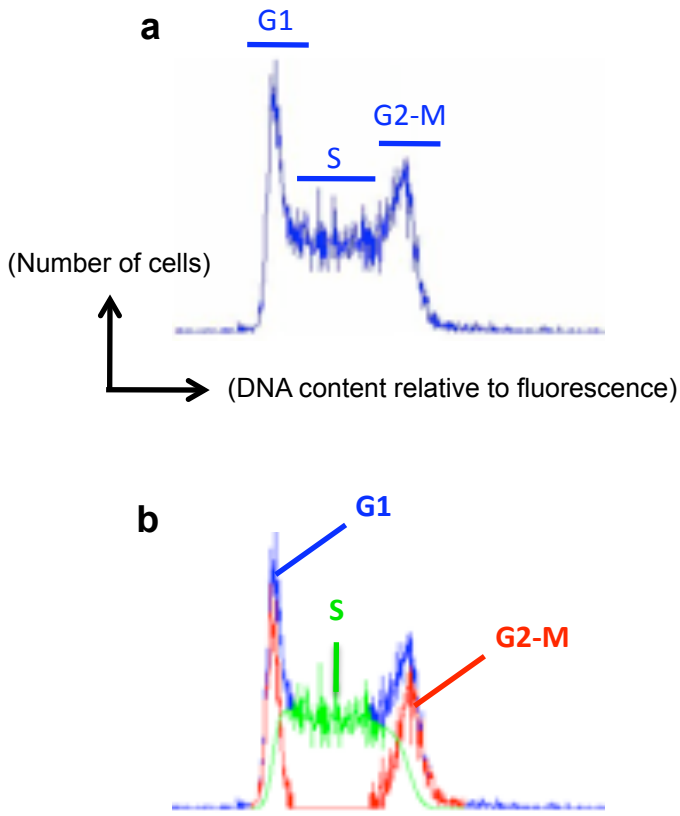
Another assay was performed to assess cell viability, the CellTiter96® Aq_{ueous} One Solution Cell Proliferation Assay (or MTS assay; Promega). This calorimetric assay was used to determine the number of viable cells in proliferation. Briefly, the provided solution contains a MTS tetrazolium compound that is bioreduced by metabolically active cells into a colored formazan product. The assay was performed in 96-well plates where 10µl of the provided solution was added to 200µl of culture medium. Following 3 hours incubation at 37°C the quantity of formazan product, proportional to the number of living cells, was measured by the absorbance at 492nm. Media alone and dead cells (death induced by the addition of 10% DMSO) were used as a control.

2.2.2 Flow cytometry analysis

Flow cytometry analysis was used to count and sort cells based on the phase of the cell cycle they were going through. In order to get significant results at least 10,000 cells have to be analysed. For each experiment, we chose to count 50000 cells per condition. Cells were harvested by trypsinisation and washed 2 times in ice-cold PBS (137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.76mM KH_2PO_4, pH7.4). They were then resuspended for 30 minutes in 200µl of staining solution composed of 10µl RNAase A and 2µl of propidium iodide (100mg/ml) per 100µl of PBS. Propidium iodide (PI) is an intercalating dye that fluoresces more strongly when bound to DNA. Therefore the amount of DNA present in the cell would be proportional to the PI signal. It had to be used in excess so that all cells present in the sample were stained. PI binds all nucleic acids, so RNAase A was used to degrade RNA for the DNA only to be quantified. Using a FACScan flow cytometer (BD FACSCalibur™), cells were then separated according to PI signal and counted. Three different populations were detected corresponding to the G1, S and G2-M cell cycle phases.

The analysis was performed using *WinMDi 2.9* and *Cylchred* software. *WinMDi 2.9* is a Windows interface able to read Flow Cytometry Standard (.FCS) files and process them to built histograms and plots. Histograms were then analysed with the *Cylchred* software that allows the deconvolution of histograms and therefore a better targeting of each cell population **(Figure 2.2)**.

Figure 2.2 Cell cycle profile using Flow Cytometry Analysis



a, DNA of untreated ES cells was stained with Propidium iodide (PI). Each cell was analysed using a flow cytometer. For each sample of about 50000 cells the cytometer generated an FCS file. Those files were read by the *WinMDi 2.9* software which then generated a corresponding **histogram**. Because the cell population was not synchronized, three sub-populations were detected based on PI fluorescence signals. Here is a typical ES cell cycle profile where most of the cells are in the S phase of the cell cycle.
b, Histogram deconvolution using the *Cylchred* software. Each sub-population was specifically targeted on the histograms and the percentage of cells was calculated.

2.2.3 Alkaline Phosphatase activity assay

Alkaline phosphatase (AP) activity is known as a marker of pluripotency (Berstine et al., 1973). The enzyme activity was detected by histochemical staining of fixed adherent cells. Briefly, medium was removed from the flask and cells were gently rinsed once in ice-cold PBS and fixed for 15 minutes in neutral formaldehyde buffer (3.7% paraformaldehyde, 112mM NaH_2PO_4, 30mM $NaH_2PO_4.H_2O$). Cells were then rinsed for 15 minutes in distilled water and incubated for 45 minutes in staining solution using Naphthol-AS-MX-phosphate (Sigma) as a substrate and Fast Red Violet LB salt (Sigma) as a coupler (0.1mg/ml Naphthol was dissolved first in 0.4% N,N-Dimethylformamide and then in 0.1M tris-HCl pH 8.3. Finally, 0.6mg/ml of Fast Red was added to the solution). Ultimately, pluripotent, AP-positive colonies appeared red.

2.3 Gene transcription analysis

2.3.1 RNA extraction and cDNA synthesis

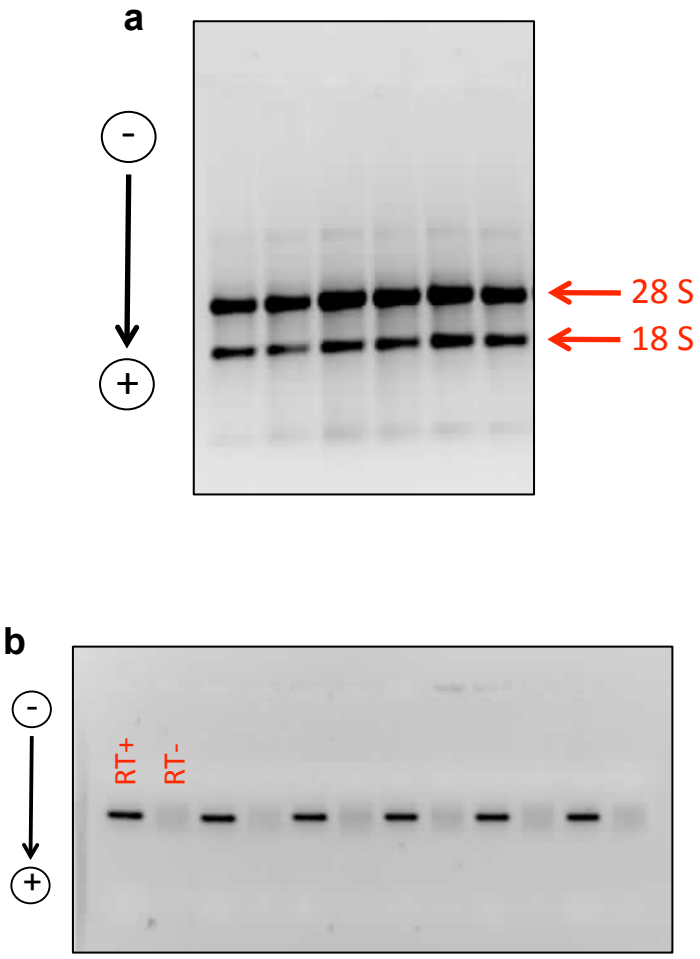
RNA was extracted using the RNeasy mini kit (Qiagen) according to the manufacturer's instructions. An on-column DNase digestion was performed during the extraction in order to yield a pure RNA fraction (RNase-Free DNase Set, Qiagen). Nanodrop analysis then provided the concentration of RNA and the 260/280 (Nucleic acid/Protein) and 260/230 (Nucleic acid/Polysaccharide) ratios per sample. Ratios, indicating the degree of nucleic acid purification, should be in a range of 1.9-2.1.

In order to check the integrity of the RNA, 1μg was loaded and run on a 1% (w/v) agarose gel.

Agarose gel electrophoresis: TAE (40mM Tris base, 1.142% (v/v) glacial Acetic acid, 1mM EDTA pH8) was used to dissolve the agarose powder and as an electrophoresis buffer. 0.5μg/ml ethidium bromide was added into the gel. After RNA electrophoresis two bands corresponding to the most abundant RNA molecules should be detected, the 18S and 28S ribosomal RNAs **(Figure 2.3.a)**.

cDNA was prepared from RNA, using the RT-Superscript-III Kit (Invitrogen) according to the manufacturer's instructions. A standard PCR was performed using β-actin primers to verify cDNA quality **(Figure 2.3.b)**.

Figure 2.3 RNA and cDNA analysis using agarose gel electrophoresis



a, **Total RNA** was extracted from ES cells and 1µg of each sample was loaded on a 1% agarose gel for electrophoresis. Two bands, corresponding to the ribosomal 18 and 28S RNAs are visible. This is because they are the most abundant form of RNA within the cell. These bands indicate the RNA integrity. A smear would have indicated a general degradation.
b, RNA was then retrotranscribed and **cDNA** amplified by standard PCR using β-actin primers. Electrophoresis on 1% agarose gel shows PCR products for RT+ and RT-. RT- was generated by retrotranscription reaction without the superscript enzyme. The absence of bands indicate there is no genomic DNA contamination.

2.3.2 Standard Polymerase chain reaction (PCR)

This technique consists in the amplification of a specific DNA fragment, or in our studies cDNA fragment, revealing the presence or absence of a specific transcript.

Primer design: Gene sequences were extracted from *UCSC genome browser* (http://genome.ucsc.edu/) and specific primers were designed using *primer3* website. Specificity of the primer set was then verified on the *Primer Blast* website (http://www.ncbi.nlm.nih.gov/tools/primer-blast/).

1µl of cDNA was added to 0.5µl of each primer (forward and reverse at 10µM) and to 20µl of PCR ReddyMix™ Master Mix (Thermo scientific). This ready-to-use master mix contains a DNA polymerase, DNTPs, salts including 1.5mM $MgCl_2$, reaction buffer and a dye facilitating the gel loading.

The PCR was then carried out by a thermocycler using the following conditions:

94°C - 5 minutes (Initialization)

94°C - 15 seconds (Denaturation)
60°C - 30 seconds (Annealing)
72°C - 30 seconds (Elongation)
} 30 Cycles

72°C - 5 minutes (Final elongation)

PCR products were analysed by 1% agarose gel electrophoresis.

2.3.3 Quantitative real-time PCR (qPCR)

qPCR is a technique based on the general standard PCR principle, but using different fluorescent dyes and detection methods with better sensitivity. We choose to use the SYBR Green dye, which absorbs blue light (λ_{max} = 488 nm) and emits green

light (λ_{max} = 522 nm). SYBR Green binds to DNA and fluoresces when intercalated in DNA, so the intensity of the signal increases proportionately to the amplification.

For each sample, 50μg of cDNA was mixed with 5μl of SYBR Mix (Qiagen), forward and reverse primers (400nM) and RNAse free water in a final volume of 10μl. Water was used as non- template control. Serial dilutions of reference cDNA (200μg, 100μg, 50μg, 25μg, 12.5μg and 6.25μg) were also analysed, allowing the construction of a standard curve.

The PCR reaction was processed by the thermocycler/detection system ABI 7900HT using the following conditions:

95°C - 10 minutes (Initialization)

94°C - 15 seconds (Denaturation)
Tm - 30 seconds (Annealing)
72°C - 30 seconds (Elongation)
} 40 Cycles

95°C - 15 seconds
60°C - 15 seconds
95°C – 15 seconds
} (Dissociation)

For every sample, a Ct (threshold cycle) value was given. A Ct is the cycle number when a threshold value of fluorescence is first detected. As a consequence, the larger the amount of targeted cDNA, the lower the Ct value **(Figure 2.4)**.

The standard curve made from serial dilutions of cDNA allowed a relative quantification. In fact, the standard curve corresponds to Ct values versus logarithm of cDNA quantity. It was also used to assess the amplification efficiency of each set of primers. A specific annealing temperature had to be established in order to get as close as possible to maximal efficiency. In optimal conditions, the curve had to show

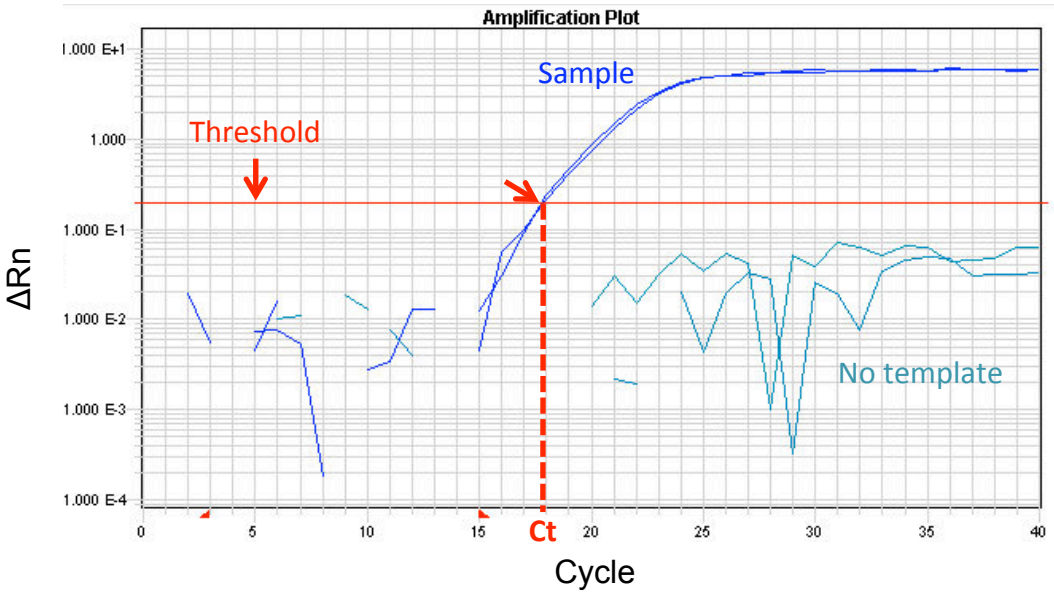
a slope of -3.3 (± 0.3), which corresponds to 100% efficiency, and a correlation coefficient closed to 1 **(Figure 2.5.a)**.

Finally, the dissociation curve reflecting the last step of the reaction, when all the double stands dissociated, had to show a single peak, indicating that a unique PCR product was generated and primer dimmers were avoided **(Figure 2.5.b)**.

β-actin or *Gapdh* were used as a house-keeper genes. Primer sets used for PCR and qPCR are listed in **Table 2.2**.

Statistical analysis was processed with the GraphPad software. Because a data set of 3 is considered a small data set, it was not possible to test the populations for normality. The choice between parametric (normal distribution) and nonparametric (population deviating from the Gaussian shape) statistical test was therefore an issue, parametric test not being robust enough and nonparametric test not being powerful enough. Nonparametric two-tailed Mann-Whitney test was usually used to test untreated and treated samples.

Figure 2.4 qPCR amplification curve

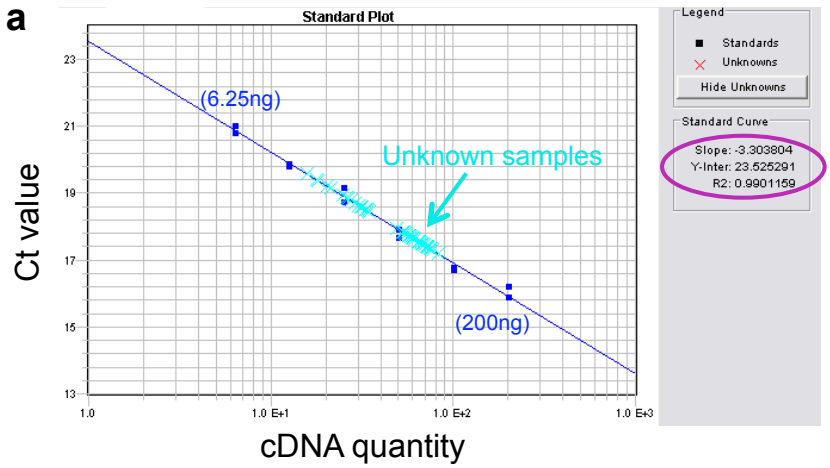


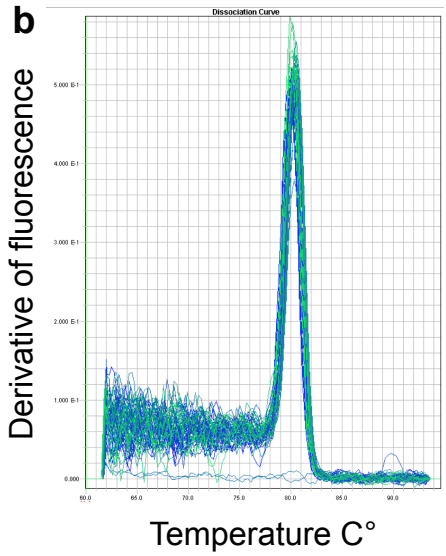
The SYBR green reporter dye was present in the PCR reaction mix. This dye binds to nucleic acids and fluoresces when trapped in the double-stranded DNA molecule. Therefore, the intensity of the florescence signal increases proportionately to the amplification.

The Passive reference Dye ROX™ was used to normalized the reporter dye signal. The fluorescence of this dye does not change during the reaction and thus, artefacts caused by changes in volumes or pipetting errors can be internally corrected. The **Rn** (normalized reporter) value equals to the fluorescence emission of the reporter dye divided by the fluorescence emission of the passive dye. The **ΔRn** value is determined by the difference between the Rn value of a sample reaction and the Rn value of a reaction that does not contain any template.

A **Ct** (threshold cycle) was determined for each sample using the amplification curve. A threshold value of fluorescence was fixed, preferentially in the exponential phase of the amplification. The Ct is the cycle number when this threshold value is first detected. Therefore, the larger the amount of targeted cDNA the lower the Ct value.

Figure 2.5 Standard curve and dissociation curve





a, A **Standard curve** was generated, using 6 serial dilutions of cDNA, from which we could extrapolate the quantity of each unknown sample. The curve also shows the amplification efficiency, and is ideal when the slope equals -3.3, and the correlation coefficient is close to 1 (good reproducibility of the reaction).
b, A **Dissociation curve** was generated at the end of the reaction when all the double stands dissociated. A single peak indicates a unique PCR product.

Table 2.2 Primers used for cDNA analysis by qPCR

	Forward primer	Reverse primer	Annealing temp (°C)
β-actin	CGTGAAAAGATGACCCAGATCA	TGGTACGACCAGAGGCATACAG	58
Gapdh	CGGCCGCATCTTCTTGTGCA	AGTGGGGTCTCGCTCCTGGA	54
Oct4/pouf5	GAGGAGTCCCAGGACATGAA	AGATGGTGGTCTGGCTGAAC	60
nanog	CTCATCAATGCCTGCAGTTTTTCA	CTCCTCAGGGCCCTTGTCAGC	62
klf4	CCAGACCAGATGCAGTCACAA	TGGCATGAGCTCTTGATAATGG	62
c-Myc	AGCAGCTCGAATTTCTTCCA	AGAGCTCCTCGAGCTGTTTG	60
Sox2	GCGGCAACCAGAAGAACAG	GTTGTGCATCTTGGGGTTCT	58
H1f0	ACCATGACCGAGAACTCCACCTC	GTGGTCACTAGGCGCTTGAT	58
Egr1	GAGCGAACAACCCTATGAGC	AGGCCACTGACTAGGCTGAA	60
Ndrg4	CCTACCTCCAGACCGAGTGA	GAAGTCCATAAGGCGTCTCG	60
Fsd1	AAGGCTAGTTGGCTGCTGTC	TCAGCAACATCTGCTTCAGG	62
Dnmt3	GGCAATTCCTTCTCTGAAGC	ACGGTTCTCCTCCTGTTCCT	60
Pgk1	AATGTCGCTTTCCAACAAGC	TTTGATGCTTGGAACAGCAG	60
Gdf3	TCAGCTTCTCCCAGACCAGGGTTTT	GCCCCGGTCCTGAACCACAG	60
Nupr1	ACAGCCTGGCCCATCCCTGT	CTCCAGCTCCGTCTCAGCGC	60
Pprc1	GGAGGCGGCGGTACAGCTCT	CTCGTCTGCCTGCCGCAACT	58
Pnrc2	GCGGCCTCCATTTTGTCGGG	ACAGCCCACGAACGAGTCACG	58
Slc6a8	GAGACTTGGACACGCCAGAT	AAAATGGGGATTCCTCCAAC	58
Anapc7	GAGTACCGCAATGCTGTGAG	TCGAGTACAGCAATGGCATC	55
Hoxb1	CCATATCCTCCGCCGCAG	CGGACTGGTCAGAGGCATC	58
Hoxb2	CGGCGCCTCCACCCTTCAGAGACC	CTTTCGGTGAGGTCCAGCAAGGC	60
Hoxb3	CAACTCCACCCTCACCAAA	GCCACCACCACAACCTTC	58
Hoxb4	CTGGATGCGCAAAGTTCAC	TCCTTCTCCAACTCCAGGAC	62
Hoxb5	CCTCTGAGCCCGAGGAAG	CCAGGGTCTGGTAGCGAGTA	60
Hoxb6	GAGACCGAGGAGCAGAAGTG	ACTGAGCTGAGACGCACTGA	62
Hoxb7	AACCGAGTTCCTTCAACATG	CGAGTCAGGTAGCGATTGTA	55
Hoxb8	TTCTACGGCTACGACCCTCT	CGTGCGATACCTCGATCCTC	60
Hoxb9	CAGGGAGGCTGTCCTGTCTAATC	CTTCTCTAGCTCCAGCGTCTGG	58
Eed	TGATTGTGTGCGATGGTTAGGCG	GCCCAATGCAAGCATCTTTTGCCA	60
Rnf2/Ring1b	CTGGGGCAGGAGCCGAAATGTC	TCCGCGCAAAACCGATGTAAACAC	60

Each set of primer was tested using samples of various known concentration of cDNA. Standard curves were built in order to assess the amplification efficiency. Different annealing temperatures were tested until the slope of the standard curve was equal to -3.3 (± 0.3) and a good correlation coefficient was obtained.

2.3.4 Microarrays

Microarray technology is a high-throughput technique used to assess the expression activity of a large number of genes under different conditions. We choose to work with the NIA 15K mouse cDNA array. One slide contains 15,247 cDNA clones derived from early embryonic cDNA libraries with a average insert size of 1.5 kb (Kargul et al., 2001).

The slides were printed by Antony Jones (College of Life and Environmental Sciences, University of Birmingham).

2.3.4.1 slide labelling

cDNA samples that had to be compared were labelled with two different fluorochromes, Cy3 and Cy5 (Amersham) using Bioprime Labelling kit (Invitrogen). Following purification of the probes (PCR purification Kit; Qiagen), quantification was performed on a spectrophotometer. Cy3 emits green fluorescence when excited at 550nm while Cy5 appears red at 650nm.

The quantity of labelled probes was obtained using the following formulas:

pmol of Cy3 probes = (Absorbance at 550nm × volume)/0.15

pmol of Cy5 probes = (Absorbance at 650nm × volume)/0.25

80 pmol of probes labelled with Cy3 and Cy5 respectively were mixed together and precipitated with 1/17th volume of 5M NaCl, 2 volumes of cold absolute ethanol and 1µl of Linear acrylamide (Ambion). Following incubation at -80ºC for 15 minutes and at 25ºC for 1 minute, the mix was spun at 11000g/4ºC for 20 minutes. The pellet was then mixed with 2 volumes of 70% ethanol and spun down again at 11000g/4ºC for 6

minutes. DNA was dried for 15-30 minutes and resuspended in 17µl of double distilled water (ddH_2O).

In the meantime, microarray slides were treated using the Pronto Backgrownd Reduction Kit (Corning) and then incubated for 2 hours at 42ºC in prewarmed prehybridisation Buffer (5XSSC, 25%(v/v)formamide, 0.1%SDS, 1%BSA). The slides were washed in 0.2XSSC (1min) and in 0.05XSSC (30sec, 2 times) and then dipped 10 times in ddH_2O and 10 times in absolute ethanol. Finally, slides were dried by centrifugation at 290g for 5 minutes.

Labelled cDNA (17µl) was mixed with 16µl of formamide, 0.64µl of 10%(w/v)SDS, 16µl of 20XSSC buffer (3M NaCl, 300mM $C_6H_5Na_3O_7$, pH7), 8µg of mouse Cot-1 DNA (Invitrogen) to prevent non-specific DNA binding and 6.4µg of Oligo PolydA (Invitrogen) to block free primers. The mixture was denatured at 95ºC for 2 minutes and left to cool at room temperature. It was then applied to the slide in a incubation chamber. The chamber was wrapped in wet towels, covered with foil and incubated for 20 hours at 42ºC. Slides were finally washed using incubation and shaking in 2X SSC/0.1%SDS (42ºC, 2min), then in 0.2XSSC (2 min) and lastly in 0.05XSSC (2 min, 4 times). Slides were dried by centrifugation for 5 minutes at 290g.

2.3.4.2 Normalization of the data

Slides were scanned using GenePix 4000A scanner: A laser excited each spot and the fluorescent emission gathered through a photo-multiplicator (PMT) coupled to a confocal microscope. PMT settings were set in order to balance the overall signal in the Cy3 and Cy5 channels. An image was obtained composed of spots going from green to red passing through the yellow colour (DNA in equal amount from both

condition). Spot alignment was made automatically and then manually using GenePix software **(Figure 2.6)**. Data were exported from GenePix, giving the mean and the median values for each wavelength (635 and 532nm) per spot. Data were then normalized using the DNMAD web-software, providing M and A values:

$$\textbf{M-value} = \textbf{Log}_2(\textbf{R}/\textbf{G}) \textbf{ and A-value} = \textbf{Log}_2(\textbf{R}+\textbf{G})/2$$

M-values correspond to the ratio between the two signals red and green and are therefore representative of differential expression between the two samples. An M-value of 1 means that the gene has been up-regulated by a 2-fold change while an M-value of -1 means that the gene has been down-regulated by a 2-fold change. A-values represent the intensity of the signals, which tells us if the gene is more or less expressed.

A box plot of M-values allows the visualization of the signal variation within an array. Each box plots show the data for an array divided in 4 groups around the median value. Because the software assumes that the expression of most of the genes is unchanged between the two conditions, median M-values are all normalized around 0 within a slide. Slides that were compared had then to be normalized at the same time in order to adjust all together **(Figure 2.7.a)**.

The red (Cy5) and the green (Cy3) dyes differ in their physical properties such as heat and light sensitivity. Therefore, there is a difference in the dye incorporation efficiency during the hybridization. As a result, an intensity-dependant normalization was needed. In a perfect situation, MA-plots should be evenly distributed around 0 across all intensity A-values. But most of the time, and as could be observed before normalization, the plot was curvy due to this difference in dye incorporation efficiency. The LOWESS (locally weighted linear regression analysis) statistical

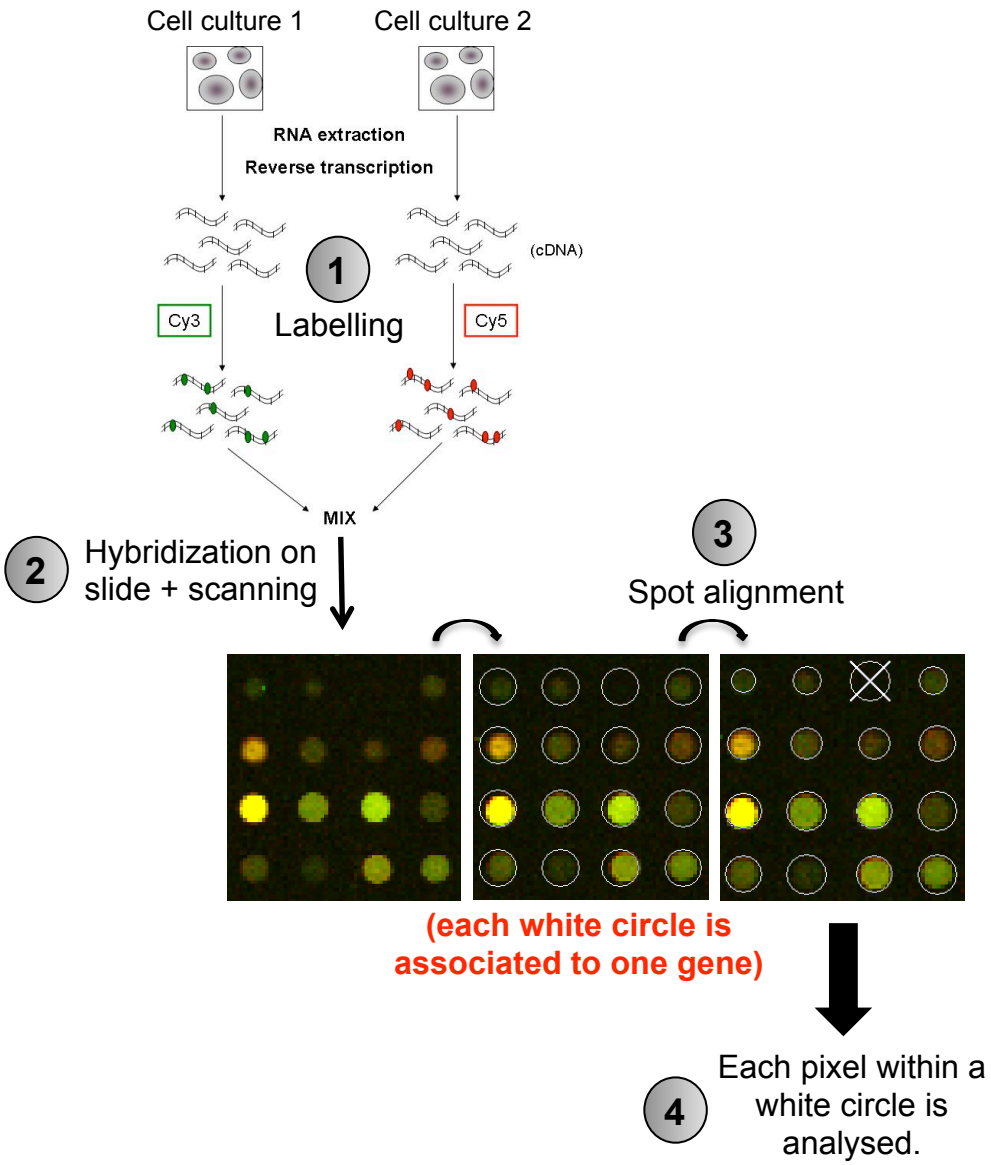
regression method was used to rectify this. In addition, it reduced the standard deviation of the data set **(Figure 2.7.b)**.

Finally, to verify the quality of the slides, distributions of red and green signal data (RG densities) were provided as well as images of the slides background and foreground. RG densities had to matched between slides and within a slide. Slide backgrounds could vary between slides but had to be low enough to be corrected. Slide foregrounds gave then a good idea of the background correction efficiency **(Figure 2.7.c)**.

2.3.4.3 Analysis of the data

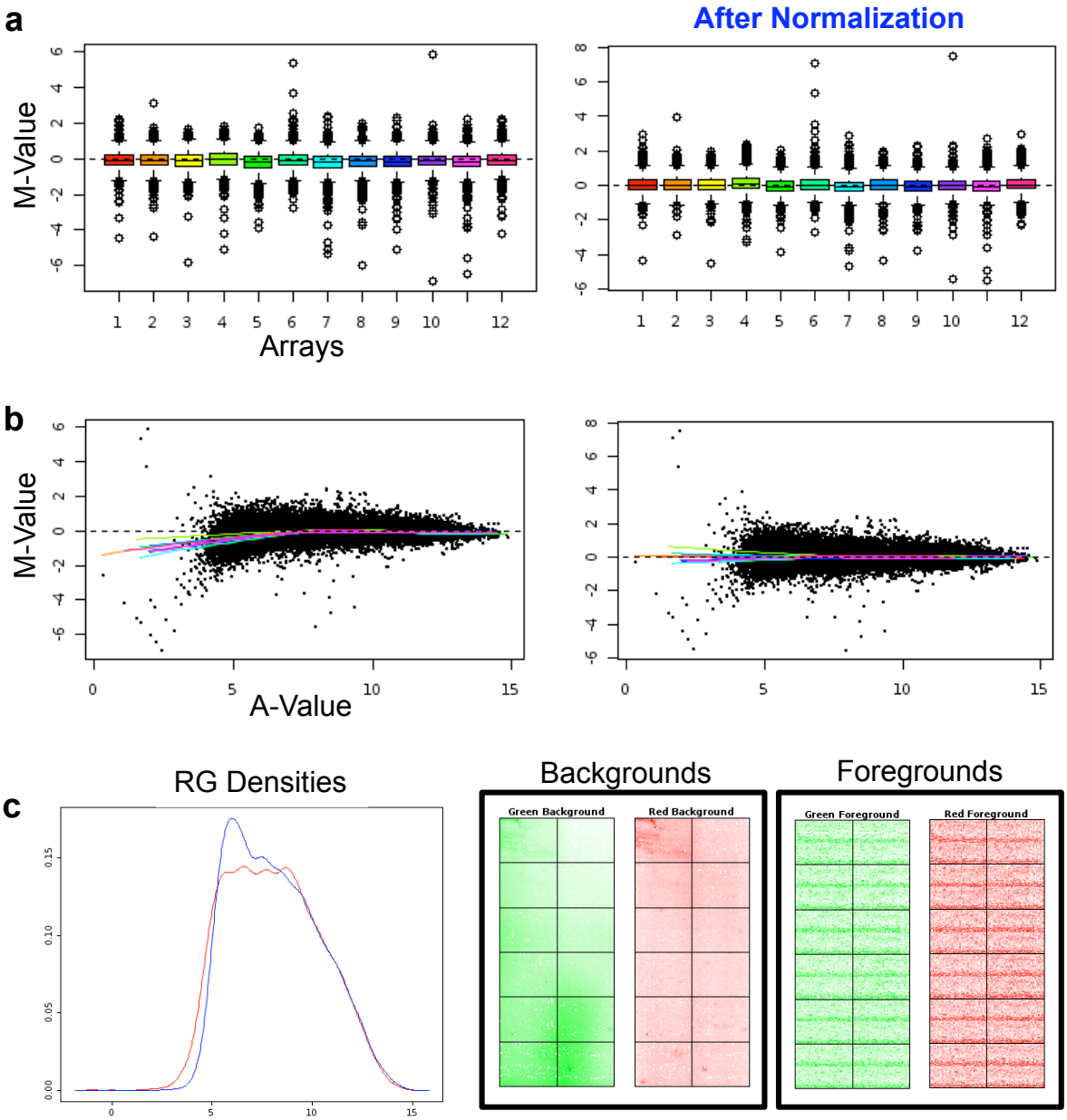
After normalization, t-tests were performed on total M-Values using biological triplicates for each condition. T-tests were performed using the web-software TMEV, a microarray data analysis tool incorporating algorithms for clustering, visualization, classification and statistical analysis. Standard p-values were then corrected by the Benjamini-Hochberg false discovery rate (FDR) correction method. The FDR correction was performed using the statistical environment “R” and data for genes with associated FDR less than 10% were considered statistically significant and selected for further analysis. For example, clustering analysis was performed, using the Web-software “DAVID bioinformatic database” or the Ingenuity and Genego softwares. Those tools, using high throughtput studies, allowed the construction of a list of genes based for example on functional similarity.

Figure 2.6 Microarray experiment protocol



Cy3 and Cy5 cDNA-dyes were used to label the two samples to be compared. Labelled cDNA were then mixed and hybridized on a NIA 15K microarray slide. Slide was scanned and an image obtained showing spots going from green to red. Each spot is associated with one gene and a yellow spot means cDNA in equal amount are present in both samples. An image composed by white circles was overlaid on the microarrays image. Each circle matches with one gene and was manually aligned with the corresponding spot. Fluorescence signals of each pixel in the circles were then analysed.

Figure 2.7 Normalization of microarray data



a, **Adjustment of M-value**. Each box plot presents the data for an array divided in 4 groups around the median value. Here M-values are all normalized around 0 within a slide.

b, **MA-plots correction**. Because the two dyes differ in their physical properties, their incorporation efficiencies vary as well. This phenomenon induces the MA-plot to be curved. A statistical regression method was used to rectify this.

c, **RG densities distribution and background correction**. In order to verify the quality of the slides, distributions of red and green signal data were provided as well as images of the slides background and foreground. RG densities have to match and backgrounds have to be low enough to be corrected.

2.4 Protein extraction

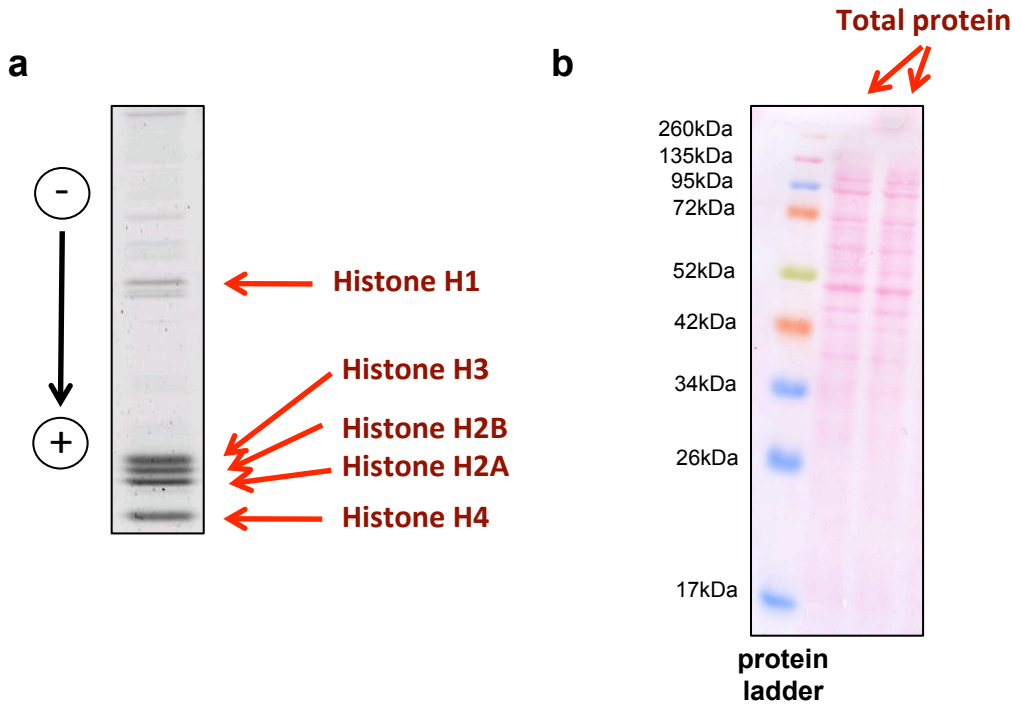
2.4.1 Histone acid extraction

Cells were harvested by trypsinisation, centrifuged (250g; 10mins) and the pellet was washed 3 times in ice-cold-PBS/5mM NA-Butyrate to prevent histone deacetylation. Cells were suspended in Triton extraction buffer (TEB; in PBS, 4mM Na-Butyrate, 0.5% (v/v) Triton X-100, 0.02% (v/v) Na-azide, 0.2mM PMSF) at a concentration of 10^7 cells per ml and left to lyse on ice for 10 minutes. Cells were pelleted at 805g for 10mins and resuspended in half of the previous volume of TEB before being pelleted again. Finally, histones were extracted by addition of 0.4M HCl at the concentration of 50µl/2×10^6 cells. After incubation on ice for three hours, the suspension was spun down and the supernatant – the histone containing fraction – was collected and kept at -20°C for further experiments **(Figure 2.8.a)**.

2.4.2 Total protein extraction

Cells were harvested by trypsinisation, washed 2 times in ice-cold PBS and suspended in lysis buffer (20mM Tris-HCl pH7.4, 10mM EDTA, 100mM NaCl, 1% Triton X-100, 1mM NaF, 1mM β-glycerophosphate, 1mM EGTA, 5mM sodium pyrophosphate and 0.1mM PMSF) at 100µl per 10^6 cells. Cells were left to lyse on ice for 30 minutes and then centrifuged at 11000g/4°C for 10 minutes. The supernatant – the protein containing fraction - was collected and kept at -20°C for further experiments **(Figure 2.8.b)**.

Figure 2.8 Protein extraction



a, Histones were extracted from ES cells and then separated according to size using SDS-PAGE electrophoresis. The gel was finally stained with Coomassie blue for the proteins to be revealed. Five main histone proteins were clearly visible.
b, Total proteins were extracted from ES cells and separated according to size using SDS-PAGE electrophoresis. Proteins were transferred to a nitrocellulose membrane and stained with Ponceau S. The Spectra™ Multicolor Broad Range Protein Ladder (Fermentas) was used for the approximate sizing of proteins.

2.4.3 Assessing protein concentration

In order to obtain protein concentrations we used the Coomassie Plus protein assay (Thermo Scientific). This assay is based on an absorbance shift from 465nm to 595nm that occurs when the reagent binds to proteins in an acidic solution, resulting in a color change from brown to blue. 20μl of protein samples were mixed with 1ml of the ready-to-use solution, incubated at room temperature for 10 minutes for a full colour change to occur and finally, the OD595nm was read. Firstly, a 2 fold-dilution BSA standard curve going from 2mg/ml to 0.125mg/ml was prepared from which the sample concentrations were read. A blank sample (20μl of 0.4M HCl) was used to calibrate the spectrophotometer (Ultrospec 2100 pro UV/visible spectrophotometer) before protein samples were analysed.

2.5 Western Blot Assay

2.5.1 SDS-PolyAcrylamide Gel Electrophoresis (SDS-PAGE)

Using this technique, proteins were separated according to size. The resolving gel consisted of 15% acrylamide (for histones, but this percentage varies according to the size of the studied protein), 0.4% N'N'bisacrylamide, 375mM Tris-HCl pH8.8, 0.1% sodium dodecyl sulphate (SDS) and was polymerised with 100μl of 10% (w/v) ammonium persulphate and 30μl of TEMED per 30ml of gel solution. The gel was overlaid with saturated iso-butanol to prevent evaporation and ensure a even polymerisation on the top. Iso-butanol was washed off before the addition of the stacking gel.

The stacking gel consisted of 3% acrylamide, 0.16% N'N'bisacrylamide, 125mM Tris-HCl pH 6.8, 0.1% SDS and was polymerised with 100µl of 10% ammonium persulphate and 10µl of TEMED per 10ml of gel solution.

5-20µg of protein were mixed with 12.5% (v/v) glycerol, 10µl of 10X Standard dissociation buffer (1M Tris-HCl pH7.5, 10mM Na_2EDTA, 10%SDS, 1.4M 2-mercaptoethanol, bromophenol blue, in distilled water) and made up to 100µl final volume in distilled water. Samples were boiled for 10 min incubated on ice for 5 min prior to loading. Gels were electrophoresed in 1X SDS Reservoir Buffer (in distilled water, 0.1%SDS, 50mM Tris base, 380mM glycine) at 30mA per gel and 400V.

2.5.2 Transfer of proteins

Proteins separated by SDS-PAGE were transferred onto a nitrocellulose membrane (Hybond C+, Amersham) using a method adapted from (Towbin et al., 1979). Transfer was performed in a cassette (Bio Rad) containing the following "sandwich": 1 sponge pad, 3 pieces of Whatman® chromatography 3MM Chr paper (Sigma), gel, nitrocellulose membrane, and again 3 pieces of Whatman® papers and 1 sponge pad. The cassette was tightly closed and placed into a tank (Bio Rad) so that the nitrocellulose membrane was between the gel and the anode. The tank was filled with transfer buffer (in distilled water, 25mM Tris base, 192mM glycine, 20%methanol) and transfer performed at 300mA for 3 hours.

2.5.3 Western blot

2.5.3.1 Antibodies

Rabbit polyclonal antisera to H4K8ac, H4K5ac, H4K16ac, H3K9ac, H3K4me3 and H2BK12,15ac were raised by immunization with synthetic peptides conjugated to ovalbumin as previously described (Turner and Fellows, 1989). Specificity was assayed by inhibition ELISA for all in-house and commercial antisera and checked by western blotting. For all antisera, cross-reaction with epitopes other than that against which the antiserum was raised was insignificant. Antibodies and working dilutions are listed in **Table 2.3**.

2.5.3.2 Western blot

Presence of proteins on the membrane was checked by red staining with Ponceau S solution (Sigma). Membrane was then rinsed in TBST (TBS (20mM Tris base, 150mM NaCl, pH 7.5), 0.1% (v/v) Tween-20) and blocked for an hour using 5% powdered milk/TBST at room temperature. The membrane was incubated with the primary antibody diluted in 5% milk/TBST for 1 hour at room temperature. Following three 10 minutes-washes in a large volume of TBST, the membrane was incubated with IRDye^{TM}800CW conjugated with an anti rabbit or anti mouse IgG secondary antibody (Rockland). The IRDye^{TM}800CW is a near infrared dye that absorbs at 774nm and emits at 789nm. After an hour incubation with the secondary antibody, membranes were once again washed 3 times in TBST. Proteins were detected and quantified using the Odyssey® Infrared Imaging System (Li-COR).

2.6 Native Chromatin Immuno-Precipitation (N-ChIP)

(Figure 2.9)

2.6.1 Chromatin preparation

Cells were harvested by trypsinisation and washed 3 times in ice-cold PBS/5mM Na-butyrate (453g, 10mins, 4°C). Pellets were suspended in TBS (in distilled water, 15mM NaCl, 10mM Tris-HCl, 3mM $CaCl_2$, 2mM $MgCl_2$, 5mM Na-butyrate) to a density of $2x10^7$ cells per ml. Equal volume of 1% (v/v) Tween-40/TBS was added as well as $1/200^{th}$ volume of 0.1M PMSF and the cell suspension was stirred on ice for a hour. Nuclei were released by homogenisation, on ice, using a hand operated Dounce all-glass homogeniser with a "tight" pestle (1 stoke per ml of cell lysate, but varied according to cell type). The nuclei yield was verified by examining a small aliquot under the microscope in a standard counting chamber. Following homogenisation, nuclei were centrifuged (805g, 20mins, 4°C) and washed in 20ml 25%sucrose/0.1mMPMSF/TBS. After centrifugation (805g, 10mins, 4°C), nuclei were suspended in 5mls of Digestion buffer (0.32M sucrose, 50mM Tris-HCl (pH7.5), 4mM $MgCl_2$, 1mM $CaCl_2$, 0.1mM PMSF, 5mM Na-butyrate). The amount of chromatin was checked using spectrophotometry. The sample was centrifuged (805g, 10mins, 4°C) and the pellet resuspended in Digestion buffer to a density of 0.5mg chromatin per ml. Chromatin was digested using 50U micrococcal nuclease for 0.5mg of chromatin at 37°C for 5mins. The reaction was stopped by addition of 5mM EDTA and samples placed on ice for 5mins. Samples were centrifuged (11000g, 5mins) and supernatants (S1 fraction) kept at 4°C. Pellets were resuspended in 1ml Lysis buffer (in distilled water, 1mM Tris-HCl (pH7.5), 0.2mM Na_2EDTA, 0.2mM PMSF, 5mM Na-butyrate) and dialysed overnight at 4°C against 2 litres of the same buffer.

Dialysed chromatin was pelleted (805g, 10mins, 4°C). The supernatant was kept on ice (S2 fraction) and the pellet was resuspended in 200μl of lysis buffer (P fraction). As previously, the amount of chromatin in each fraction was checked using spectrophotometry. Samples (in water, 2μg chromatin, 0.1%SDS, 2μl loading buffer, 25μl final volume) were analysed by 1.2% agarose gel electrophoresis. It was important not to place ethidium bromide in the gel due to the presence of SDS in the samples. The gel was stained afterwards in electrophoresis buffer containing ethidium bromide **(Figure 2.10.a)**. S1 and S2 fractions were combined and chromatin concentration was measured **(Figure 2.10.b)**.

2.6.2 Immuno-precipitation

For immunoprecipitation of chromatin, SAS cut or affinity purified antibodies were used. According to the antibody, 50-100μl was added to 100 μg of unfixed chromatin (S1+S2) **(Table 2.3)**. The final volume was made up to 1ml with Incubation buffer (in distilled water, 75mM NaCl, 20mM Tris-HCl, 20mM Na-butyrate, 5mM EDTA, 0.1mM PMSF). After overnight incubation at 4°C on a slow turntable, 200μl of 50% (w/v) Protein A Sepharose^{TM} Cl-4B (Amersham Biosciences) were added and incubation continued for a further 3 hours at room temperature on a fast turntable. The sample was centrifuged (11000g, 10mins) and supernatant was kept as the UNBOUND fraction (UB). The pellet was resuspended in 1ml of Wash buffer A (50mM Tris-HCl (pH7.5), 10mM EDTA, 5mM Na-Butyrate, 50mM NaCl) and layered onto 9mls of the same buffer using a siliconised pasteur pipette in a 15ml siliconised polypropylene centrifuge tube. After centrigugation (805g, 10mins), the pellet was sequentially washed with 10mls of Wash buffer B (50mM Tris-HCl (pH7.5), 10mM EDTA, 5mM

Na-butyrate, 100mM NaCl) and 10mls of Wash buffer C (50mM Tris-HCl (pH7.5), 10mM EDTA, 5mM Na-butyrate, 150mM NaCl) and finally resuspended in 1ml of Wash buffer C. Following transfer in siliconised eppendorfs, the sample was centrifuged (11000g, 10mins) and the pellet resuspended in 250μl of 1%SDS/incubation buffer for 15mins at room temperature on a fast turntable. This step eluted the bound material from the Protein A Sepharose, so after centrifugation (11000g, 10mins) the supernatant was kept as the Bound 1 fraction. The pellet was washed again in 250μl of 1%SDS/incubation buffer and centrifuged (11000g, 10mins) and the supernatant mixed with the bound 1 fraction to obtain 500μl of BOUND fraction (B). In order to reduce the SDS concentration to 0.5%, 500μl of incubation buffer was added to the bound fraction.

2.6.3 DNA isolation

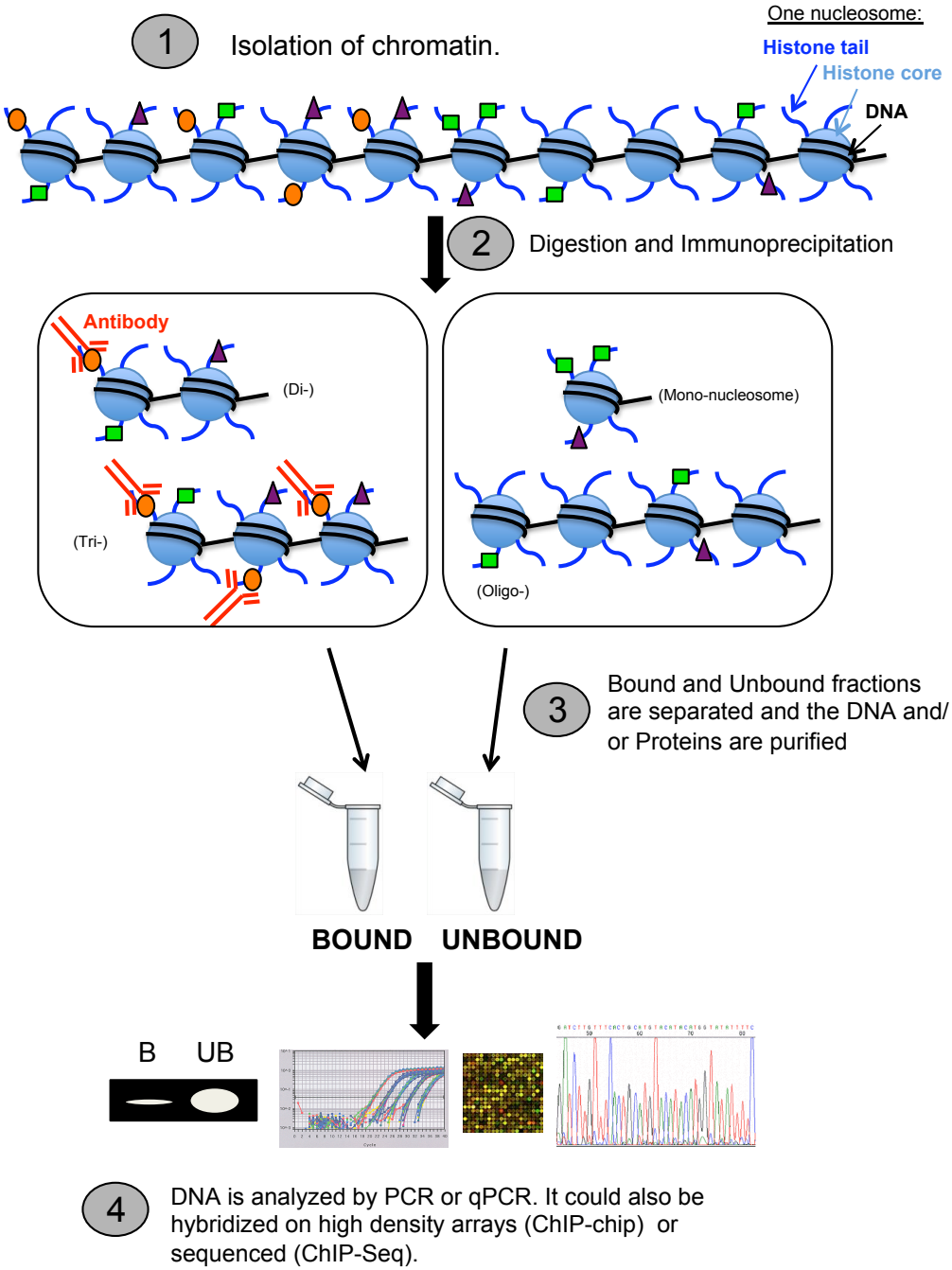
Each fraction was added to 333μl of phenol/chloroform, then vortexed and centrifuged (805g, 10mins). The Upper phase was added to 1ml of phenol/chloroform, then vortexed and centrifuged (805g, 10mins), and then re-extracted with two 1ml volumes of chloroform. Finally, the upper-phase was mixed with 100μl of 4M LiCl, 2.5μl of glycogen (20mg/ml) and 4mls of ice-cold ethanol. Following overnight incubation at -20°C in order to precipitate the DNA, the sample was centrifuged (1811g, 20mins, 4°C), The pellet dried, and resuspended in 250μl of distilled water.

DNA samples were initially analysed on 1% agarose gels **(Figure 2.10.c)**. DNA in bound and unbound fractions was then quantified using Quant-it PicoGreen (PG) Reagent (Invitrogen). 2μl of DNA was added to 48μl of TE buffer (10mM Tris, 1mM

EDTA, pH8) and 50μl of PG (1/200 in TE buffer). The assay was performed in a 96-well plate using the Wallac Victor^3 1420 fluorometer. The percentage pull down was finally calculated (=(B/UB+B)*100) and the UB was diluted to the B in distilled water to ensure that equal amounts of DNA were analysed.

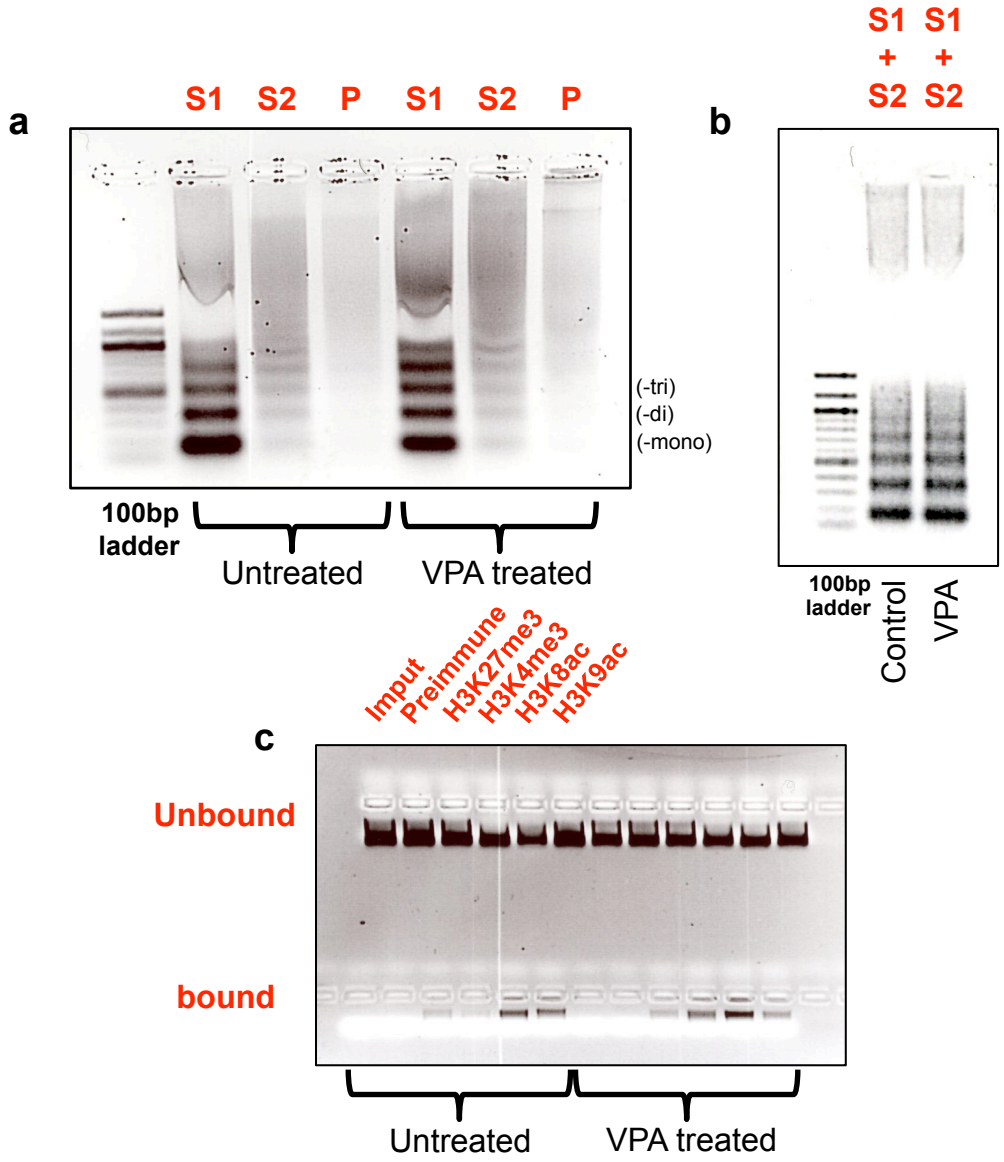
In order to detect area of mark enrichment, specific regions of specific genes were analysed by qPCR **(Table 2.4; 2.5)**. The ratio (B/UB) was calculated and considered as enrichment if equal or above 1.

Figure 2.9 Native chromatin immunoprecipitation



Nuclei were isolated and chromatin was extracted from ES cells. Chromatin was then digested with the micrococcal nuclease, generating a ladder of ≈200bp fragments. Chromatin was immunoprecipitated using antibodies raised against specific post-translational histone modifications. Protein A sepharose was used to separate bound (B) and unbound (UB) fractions. DNA was then purified by phenol-chloroform extraction. Analysis of the mark enrichment was done at specific sites using PCR or qPCR. High-throughput technologies such as hybridization on high density arrays or sequencing could be used for global studies.

Figure 2.10 ChIP analysis using gel electrophoresis



a, Chromatin was extracted from ES cells untreated or treated with VPA (1mM, 8hrs). Chromatin was then digested with the micrococcal nuclease. S1 and S2 fractions correspond to the soluble digested chromatin, respectively after digestion and after overnight dialysis of the pellet. P fraction corresponds to the pellet resuspended in lysis buffer after overnight dialysis.
b, Good quality S1 and S2 fractions were pooled for immunoprecipitation. The imput chromatin was visualised using 1.2% agarose gel electrophoresis.
c, After immunoprecipitation, DNA from unbound and bound fractions was isolated using phenol-chloroform extraction. DNA was then analysed on a 1.2% agarose gel. As expected, more genomic material was present in the unbound fractions. Also, no DNA was detected in bound fractions of samples which were immunoprecipitated with no antibody (imput) or preimmune serum.

Table 2.3 Antibodies and dilutions

Antibody	Raised against	Dilution for Western blot	Dilution for ChIP (for 100µg of chromatin)	Dilution for immuno-fluorescence
R403	H4K8ac	1/1000	100µl	
R41	H4K5ac	1/500		
R252	H4K16ac	1/1000		
R607	H3K9ac	1/500	50µl	
R612	H3K4me3	1/400	50µl	
Millipore 07-449	H3K27me3	1/2500	5µl	1/400
Ab-4729	H3K27ac	1/500		1/200
R209	H2BK12,15ac	1/400		
Ab-1791	H3 C-term	1/10000		
Ab-19857	Oct4	1/500		
Ab-26648	Klf4	2.5µg/ml		
Ab-8227	Actin	1/2500		
Rockland 611-131-122	Rabbit IgG	1/1000		
Rockland 611-131-121	Mouse IgG	1/1000		
Sigma F1262-1ML	Rabbit IgG			1/50

Rx = Rabbit house antibodies; Ab-x = Abcam antibodies; H = histone; K = Lysine; ac = acetylated; me3 = tri-methylated

Table 2.4 Primers used for ChIP-DNA analysis by qPCR

	Forward primer	Reverse primer	Annealing temp (°C)
Egr1/1	GGCAAGCTGGGAACTCCA	ACCCGGAGTGACGTGAAG	60
Egr1/2	GGCTTCCTGCTTCCCATA	TGGGATCTCTCGCGACTC	60
Egr1/3	CTGCACCCCGCATGTAAC	GCGAGCTGGAGAACTGATGT	60
Egr1/4	TGTTGTTTCTCCTGGGCTTG	AGTTCCCAGAACCACCACAC	60
H1f0/1	GCCCTTCCAACTCTTTCAGG	CGGCCCCAACTAACACAC	60
H1f0/2	GAGTGGGGCTCCTGGTAGTC	GGCACATTTCCCCTTTGC	60
H1f0/3	GGTGCGTTTTTCTTTAACTGG	TCTCCCGAAGACCCACCTAC	60
Ndrg4/1	AGCTTCTCCTCGGGTCTAGC	GCACCATCCATTAGGTACAGC	60
Ndrg4/2	CTTGAGCGGCAGAAGCTG	CCTCCCTGCAAGGTTTCTG	60
Ndrg4/3	GAGCTCTACCCCACCACAGA	GACCCTCACTCACCCCATC	60
Ndrg4/4	CCTATCACAGGGAAGGCTCA	CAGCCTGGGCTACAGAAAGA	60
Anapc7/1	TTAGCGCAGCAAAACTACGA	ACGGACGTGCCTGAGTAAAA	60
Anapc7/2	GAGCCTCTGCTCTCGGTTC	CGGGTTGTTGTTACTCATGG	60
Anapc7/3	GCCTGGCTCTTCACTGAGAT	CAACCTCTCACATTCCTAGCAC	60
Slc6a8/1	CAACCTTGGTGAAGGATGGT	CATTCGTTCCAGGATCACCT	60
Slc6a8/2	GAGAGGGACAGTTGGAGTGC	CAAGTGGCTGTCCTGTTCG	60
Slc6a8/3	TGGCTCCTGTTTCTGCTTCT	GAGGCTTCCTTTCCTGCTCT	60
Nanog/1	GCAGCCGTGGTTAAAAGATG	AGGGAGGGAAAGCTTAGGG	60
Nanog/2	ATCCACTGAGCCATCTCACC	CTGGTCTGCAGAGCTAGTTCAA	60

Each set of primer was tested using samples of various known concentration of cDNA. Standard curves were built in order to assess the amplification efficiency. Different annealing temperatures were tested until the slope of the standard curve was equal to -3.3 (± 0.3) and a good correlation coefficient was obtained.

Table 2.5 Primers used for ChIP-DNA analysis by qPCR (2)

	Forward primer	Reverse primer	Annealing temp (°C)
Hoxb1/1	AGCCCTACAGCCTTGGGGTGG	TTCAACCCAGCTGCCCCTCAC	58
Hoxb1/2	TCCTAGGGGATCGCTGGCGG	GCCGTTTGGGGGAGACACCC	58
Hoxb1/3	CCGCAGCCCCCATACGGAAC	CGGGGAGTGAGAGTGCTGGGT	60
Hoxb2/1	TCCCCCAAGGTATGGAAGCCTCT	CCTACCACCACCTCAGCCCCC	60
Hoxb2/3	CAAGCCGGAGATGGGCCTGC	GAGGCGGCTGGAGAAGGGGA	60
Hoxb4/1	GCTACACCCGCAGCCCCAAC	GAGACCCCCTCCTCGCCTGG	60
Hoxb4/2	GGTGGCCATTGGCTCGGAGG	CGCACGGAGGGAACTTGGGG	60
Hoxb4/3	TGTCCCCTCGGGCTCCAGTG	GGGCTCTTTGCACGCGGAGT	60
Hoxb5/1	AGAGGGGATAGCCGCACCCT	CTGGCTCTGTGCAACCGCCA	58
Hoxb5/2	TGGCCGCATACATAGCAAAACGA	GGAGGGCTTGATTTGTGGATCGTGG	58
Hoxb5/3	GCCCTGCACTAACGGCGACA	TCCTCGGGCTCAGAGGACGC	60
Hoxb7/1	CACCAAGTTGCGAACATTCA	CAGCTTGCAGGCGAGATT	60
Hoxb7/2	ACCGACACTAAAACGTCCCCGA	AGCGAAAACCGAACTTGCGGC	58
Hoxb7/3	GCGCCAAGGAGCAGAGGGAC	CGGAGCCCTTCTGGGGACCG	58

Each set of primer was tested using samples of various known concentration of cDNA. Standard curves were built in order to assess the amplification efficiency. Different annealing temperatures were tested until the slope of the standard curve was equal to -3.3 (± 0.3) and a good correlation coefficient was obtained.

2.7 Immunofluorescent Labelling:

2.7.1 Immunofluorescent labelling of cells grown on coverslips

Cells were grown on a coverslip in a 24 well plate. Tissue culture medium was removed and cells gently washed three times with 1XPBS. Cells were fixed in 4%(v/v) formaldehyde/PBS for 30 minutes at room temperature before blocking with 1%BSA/0.5%Triton-X100/PBS for 30 minutes at room temperature. The primary antibody was diluted in the blocker solution **(Table 2.3)** and a minimum of 150μl was applied on cells. Following incubation at 4°C for an hour, cells were washed three times with 1XPBS. The last four steps were repeated with the secondary antibody, a goat anti rabbit-IgG labelled with fluorescein isothiocyanate (FITC, Sigma) . The coverslip was finally carefully lifted from the well using fine tweezers and placed on clean tissue, upper side facing up. 8μl of DAPI (Sigma) diluted to 1μl/ml in Vectashield mounting medium (H-1000; Vector Labs) was applied on cells and a glass slide (VWR 1.0-1.2mm, ethanol washed) gently lowered down on coverslip. Bubbles were carefully eased away and the slide and coverslip sealed together with nail varnish. Slides were kept at room temperature in the dark until images were taken using fluorescent microscopy.

2.7.2 Immunofluorescent labelling of metaphase chromosomes

In order to block cells in metaphase, 10μl/ml of Colcemid (Gibco, 10μg/ml) was added into the culture medium for 3 hours. Cells were harvested by trypsinisation and the suspension centrifuged at 250g for 5 minutes. The pellet was washed twice with ice cold PBS before being resuspended to $2x10^5$ cells/ml in 75mM KCl for a 10 minutes incubation at room temperature. Chromosome spreads were then prepared

on glass slides using a Shandon Cytospin 4. $4x10^4$ cells per chamber were added and spun at 1800rpm for 10 minutes. Slides were then immersed in KCM buffer (120mM KCl, 20mM NaCl, 0.5mM EDTA, 0.1%(v/v) Triton-X-100, 10mM Tris-HCl pH8) for 10 minutes at room temperature. 50μl of primary antibody diluted in KCM/0.1%BSA **(Table 2.3)** were applied to slides and covered with Parafilm for one hour incubation at 4°C in a humid chamber. Slides were then washed twice by immersion in KCM buffer for 5 minutes. As previously, 50μl of secondary antibody, a goat anti rabbit-IgG labelled with fluorescein isothiocyanate (FITC), diluted in KCM/0.1%BSA was applied on slides and covered with Parafilm for one hour incubation at 4°C in a humid chamber. Following two 5 minute-washes with KCM buffer, slides were immersed in KCM/4%(v/v) formaldehyde for fixation. Slides were then rinsed with distilled water. Finally, 7.25μl of DAPI (Sigma) diluted to 1μl/ml in Vectashield mounting medium (H-1000; Vector Labs) was applied on cells and a glass coverslip gently lowered down on the slide and sealed with nail varnish. Slides were kept at room temperature in the dark until pictures were taken using fluorescent microscopy.

3. Results

3.1. Epigenetic inheritance in ES cells: Role of histone acetylation

3.1.1 Introduction

Many processes have been described in the literature as epigenetic events. Genomic imprinting, X-chromosome inactivation, differentiation or phenomena like reprogramming and carcinogenesis require the establishment and long-term maintenance of specific gene expression patterns. We believe that cellular transcriptomes are partly specified by epigenetic mechanisms including the post-translational modifications of histones. There is now evidence that these proteins play a major role in gene transcription, directly by changing the chromatin conformation (Shogren-Knaak et al., 2006) and indirectly by acting as binding sites for other proteins like chromatin modifiers or transcription factors (Taverna et al., 2007). Thus, for a particular phenotype to be maintained through mitosis, epigenetic modifications have to survive replication and be transmitted to daughter cells. Little is known about how epigenetic marks are transmitted during the cell cycle but some mechanisms have been described, implicating pre-existing histone modifications as templates for newly formed nucleosomes (Nightingale et al., 2006). We decided to test this idea by asking whether or not an environmentally-induced epigenetic change could be heritable through mitosis. If histone modifications are read and duplicated during the cell cycle, we hypothesized that several generations of cells might keep the memory of an epigenetic change induced by drug treatment. In order to test our hypothesis we designed a model system using mouse embryonic stem (ES) cells and the histone deacetylase inhibitor (HDACi) valproic acid (VPA). Cells were treated for different periods of time with VPA so that global histone acetylation levels were

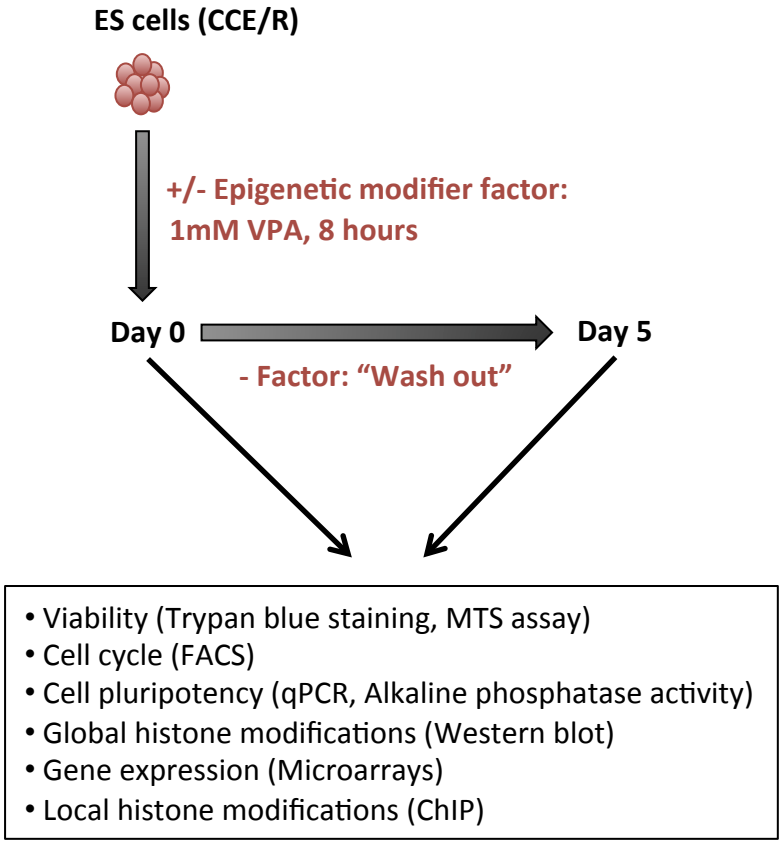
increased. Following the treatment, the effect of hyperacetylation on different parameters in our cell line was measured. In the meantime, VPA was removed from treated cells and these cells were kept in culture free from VPA for a further 5 days. After this wash out period, the same series of experiments were carried out in order to detect any persistent effect **(Figure 3.1)**.

Mechanisms that allow epigenetic inheritance remain unclear. However, it has been suggested that duplication of the marks associated with chromatin structure and DNA replication are connected (Zhang et al., 2000). Therefore, ES cells offered a good model for our study as they divide very rapidly - due to a lack of checkpoint controls, they show a significantly shortened G1 gap phase in comparison with most somatic cells. As a result, cells in an ES cell culture spend most of their time in the S phase (replication phase) of the cell cycle and display an unusually rapid doubling time of approximately 8-10 hours (Stead et al., 2002). This particular feature of ES cells constituted an advantage for our model. Valproic acid being a toxic drug, we wanted to treat the cells for relatively short periods of time, but long enough for most of the cells to go through a complete cell cycle.

In an embryo the pluripotent state is transient while in an ES cell culture this unique characteristic has to be maintained. Constantly fighting differentiation and maintaining a chromatin structure with high plasticity (Meshorer et al., 2006), ES cells are thought to be highly sensitive to their environment and were therefore more likely to respond to VPA treatment.

For those reasons, ES cells were considered a suitable model for our study.

Figure 3.1 Model system #1:
Is it possible to induce a heritable epigenetic change in ES cells?



The model system was designed to measure the effects of the HDACi Valproic acid (VPA) on ES cells (CCE/R line) and whether or not those effects could persist through the cell cycle.
ES cells were treated for 8 hours with 1mM VPA. Because VPA is a HDACi we expected the treatment to induce hyperacetylation of histones. After 8 hours, cells were harvested and tested for several parameters including cell cycle profile, histone modifications levels or global gene transcription activity. In order to detect any heritable change, a pool of cells were kept in culture for 5 days in VPA-free medium following the treatment. After the wash out, cells were harvested and tested again.

3.1.2 Effect of valproic acid treatment on viability, self-renewal and pluripotency of ES cells

The potential anti-cancer activity of Valproic Acid (VPA) has been linked to its ability to induce cell cycle arrest, apoptosis or differentiation (Bolden et al., 2006). We wanted to limit those effects in our study because we were specifically interested in the epigenetic long-term effect(s) of the drug. As ES cells are extremely sensitive to the environment, we had to choose a concentration of VPA we could use to induce hyperacetylation without dramatically disrupting the cell cycle, or the viability, or pluripotency of the cells.

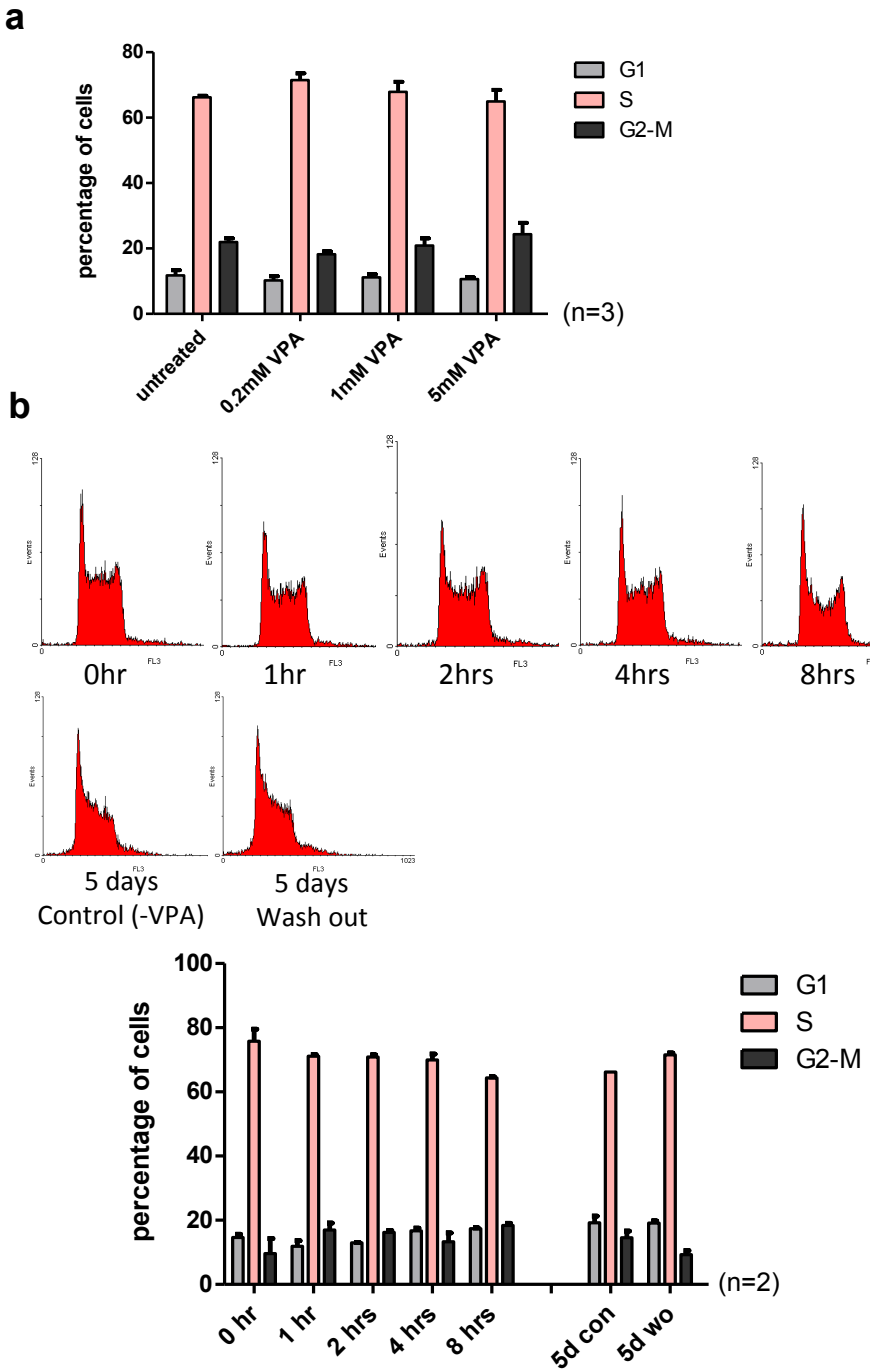
3.1.2.1 Cell cycle

As mentioned earlier, due to a shortened G1 phase, ES cells divide very rapidly in comparison to differentiated cells. In a first experiment, ES cells were treated for 8 hours with different concentrations of VPA (0.2mM, 1mM and 5mM) and the percentages of cells in each phase of the cell cycle were measured using flow cytometry. As expected in untreated cultures, more than 60% of the cells were in the S-phase of the cell cycle while a mean of 11.8% were in the G1-phase and 22% in the G2-M-phase. None of the treatments significantly changed this distribution. Indeed, the proportions of cells in S-phase stayed over 60% and proportions of cells in G1-phase and G2-M-phase remained very consistent **(Figure 3.2.a)**. As 1mM VPA appeared to be well tolerated by the cells, and (as seen later) was enough to induce hyperacetylation, we decided to use VPA at this concentration for future experiments.

A time-course experiment was then performed, which would have allowed the detection of any major effect occurring at earlier time points within 8 hours treatment. 1mM VPA treatment was carried out for 1, 2, 4 and 8 hours. Cells were harvested at each time point and the percentages of cells in each phase of the cell cycle were measured. Despite the treatment, a very high percentage of cells were still undergoing replication at each time point. In fact, the percentage of cells in S-phase never went below 60%. Although this percentage was slightly decreased during treatment (76% to 65%), we concluded that the treatment did not induce major changes in the cell cycle profile as this small drop in the number of replicating cells could simply reflect normal time-variations due to culture conditions.

In addition to the time course, we decided to look at the cell cycle profile of cells that were treated for 8 hours with 1mM VPA and washed out for 5 days. A control flask (untreated cells) was harvested at the same time. Again, more than 60% of the cells were found in S-phase in both samples. The numbers of cells in each phase showed small differences between the two samples, but this variation was not greater than that observed during the 8 hours time course **(Figure 3.2.b)**.

Figure 3.2 Effect of VPA treatment on ES cell self-renewal



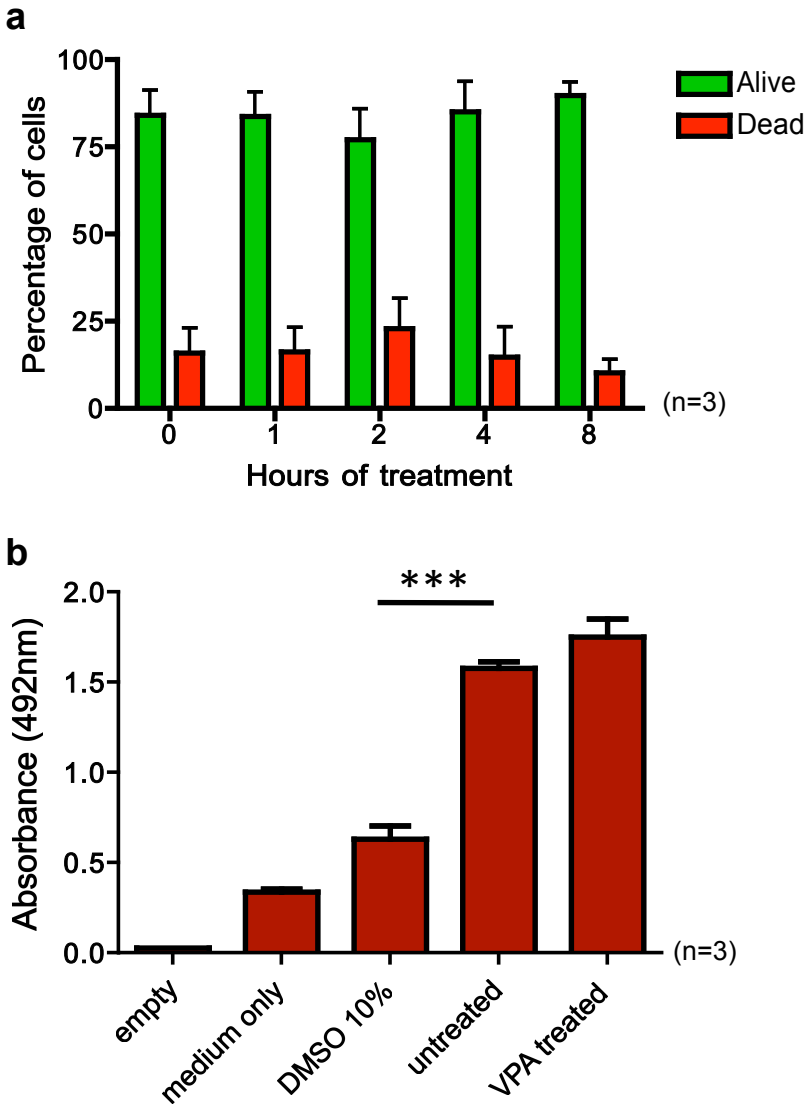
CCE/Rs were first treated for 8hrs with 0.2mM, 1mM or 5mM VPA **(a)**. In a second experiment, CCE/R were treated with 1mM VPA for 1, 2, 4 or 8 hours. Following the 8 hours treatment, VPA was removed (Wash-out, wo) from some cells for 5 days before analysis **(b)**. For each sample, 50000 cells were stained with propidium iodide (PI). The intensity of the signal from each cell was then analysed using a Flow cytometer in order to obtain the mean percentage of cells in each phases of the cell cycle. For the time course, histograms from one of the two independent experiments used to calculate percentage of cells are shown.

3.1.2.2 Cell Viability

Affecting several pathways, VPA is known to induce cell death but also to have an anti-apoptotic activity (Bielecka and Obuchowicz, 2008). Therefore, we decided to measure ES cell viability after VPA treatment. Cells were treated with 1mM VPA for 1,2,4 and 8 hours and stained with Trypan Blue. Metabolically active cells were able to exclude the dye while dead cells appeared blue. For each time point, white and blue cells were counted using a Malassez hemocytometer. On average in untreated cultures, 84.1% of the cells appeared alive and 15.9% of the cells were dead. Although cell death slightly increased after 2 hours of treatment, the proportions of alive and dead cells didn't significantly vary during the treatment **(Figure 3.3.a)**.

Because Trypan blue staining is not very accurate, cell viability was assessed a second time using a calorimetric assay that allows the quantification of metabolically active cells in proliferation. During the assay, a compound was bioreduced by living cells into a coloured product. The quantity of product measured in the medium by the absorbance at 492nm was directly proportional to the number of living cells. Untreated cells and cells treated with 1mM VPA for 8 hours were compared. Empty well, medium only and cells treated with 10% DMSO were used as negative controls. As expected, the difference between the dead cells sample (+DMSO) and untreated cells was statistically significant, while no effect was detected between untreated and VPA-treated cells **(Figure 3.3.b)**.

Figure 3.3 Effect of VPA treatment on ES cell viability



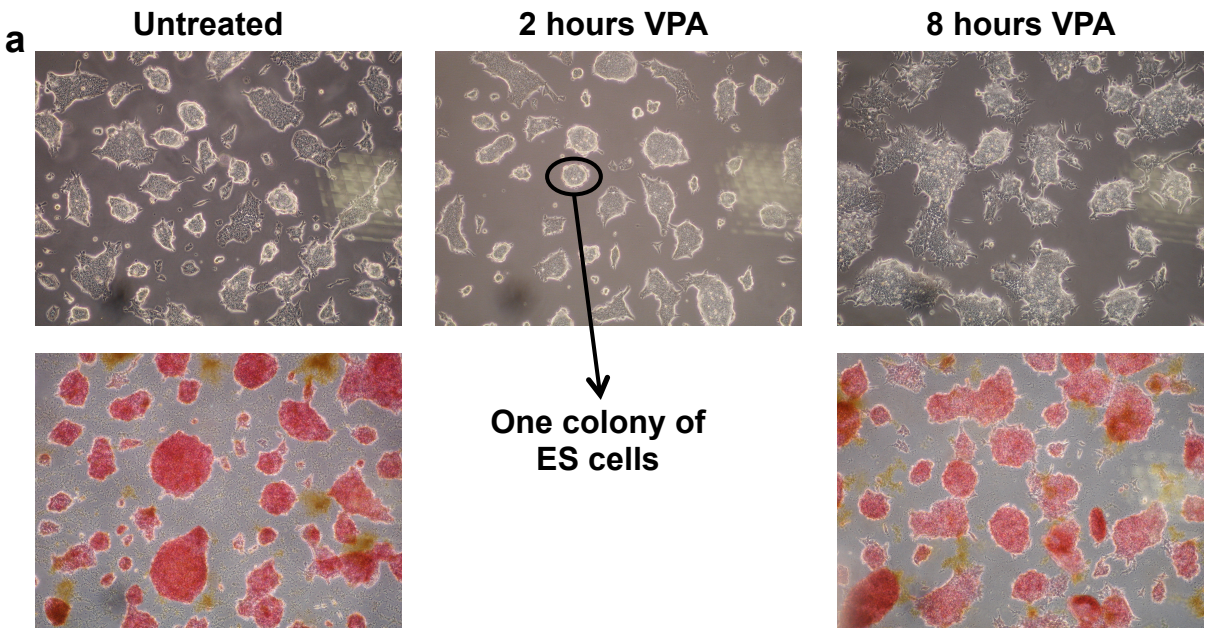
a, CCE/Rs were treated with 1mM VPA for 1,2,4 or 8 hours and stained with trypan blue. Alive and dead (remaining blue) cells were counted using a malassez hemocytometer. Means of alive and dead cell percentages were calculated at each time point (n=3).
b, CCE/Rs were treated with 1mM VPA for 8 hours or with 10% DMSO in order to induce cell death. The calorimetric MTS assay was performed to assess viability of cells in each conditions. An empty well and a well containing media only were used as additional negative controls. The quantity of the colored formazan product, directly proportional to the number of living cells, was measured by the absorbance at 492nm (n=3).
Two-tailed unpaired t-test were done using the GraphPad software.

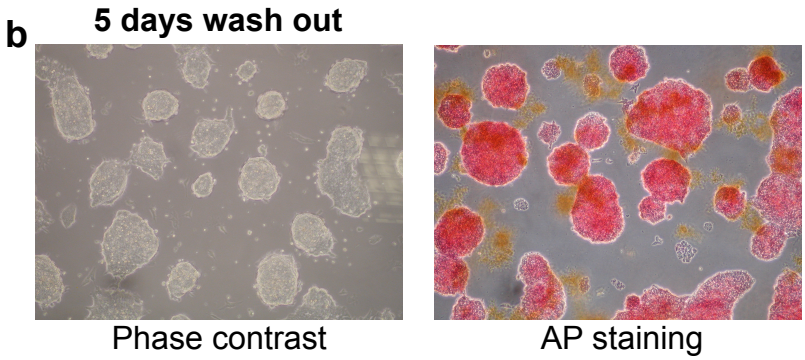
3.1.2.3 Cell pluripotency

Although 1mM VPA didn't induce massive changes in cell cycle profile or affect the viability of our ES cells, it did induce an important change in cell morphology. ES cells commonly grow in colonies where one cell often can't be distinguished from another. However, when cells were treated with 1mM VPA for 8 hours the colony morphology was affected. Single cells were becoming visible at the edges of the colonies by adopting a monolayer growth pattern. We speculate that the cytoskeleton of the cells was affected by the treatment, possibly as a consequence of an early differentiation process. Therefore, we decided to test untreated and treated cells for alkaline phosphatase activity, a marker of pluripotency (Berstine et al., 1973). As expected, colonies of untreated ES cells stained red which confirmed a strong enzymatic activity. After 1mM VPA treatment for 8 hours, the staining still appeared uniform but was noticeably weaker. At this stage, it was difficult to say if this was a consequence of an early differentiation process or if this was because cells within the colonies were less compacted and therefore the colour more diluted **(Figure 3.4.a)**.

Cells treated with 1mM VPA for 8 hours and washed out for 5 days were tested for alkaline phosphatase activity as well. The test revealed cells with a normal morphology and strong alkaline phosphatase staining **(Figure 3.4.b)**.

Figure 3.4 Effect of VPA treatment on cell morphology and Alkaline phosphatase activity





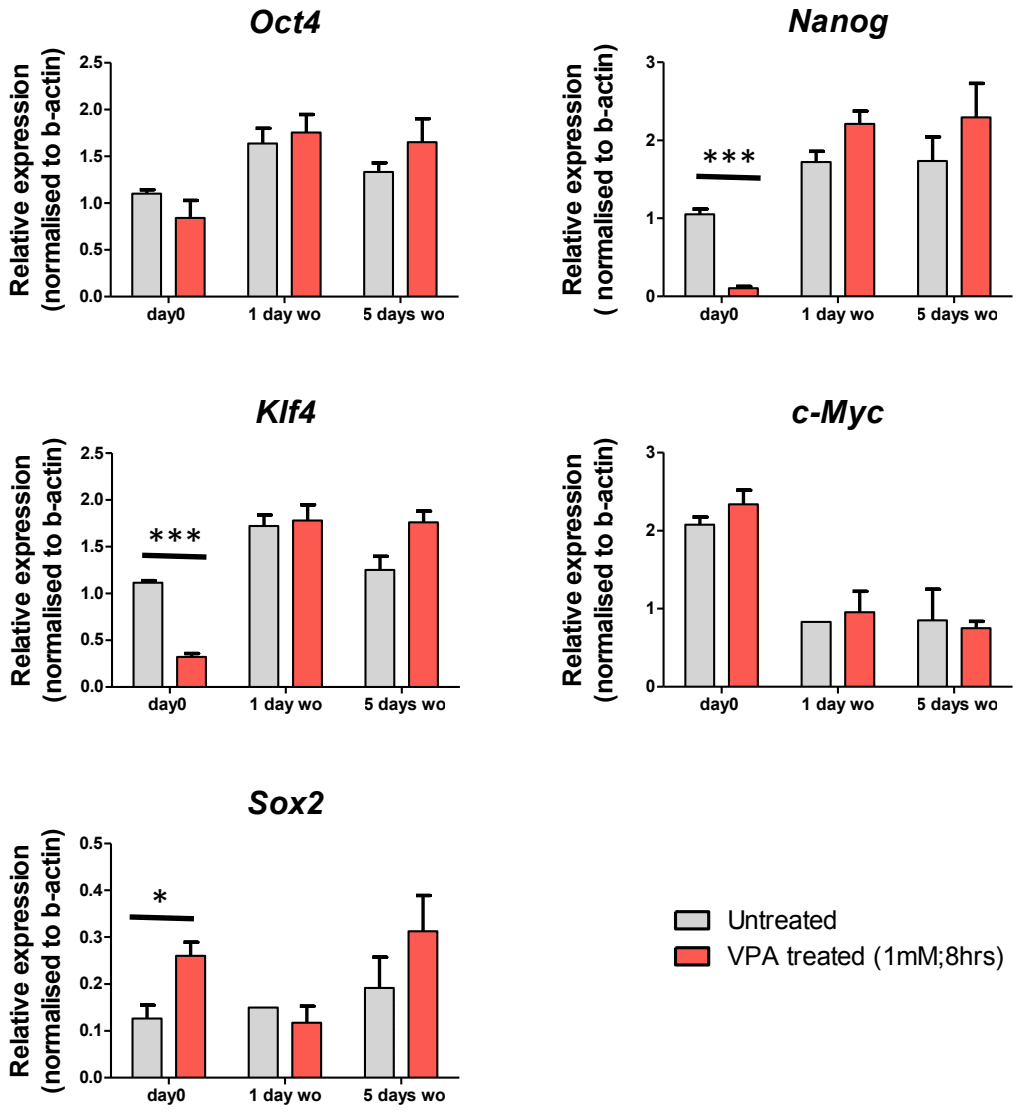
CCE/Rs were treated with 1mM VPA for 2 and 8 hours **(a)** or treated with 1mM VPA for 8 hours and washed out for 5 days **(b)**. Light microscopy was used to assess the cell morphology at each time point. In addition, untreated cells, 8 hour treated cells and washed out cells were stained for alkaline phosphatase activity (AP), a marker of pluripotency. Pluripotent, AP-positive colonies appeared red. Pictures were taken using 20X magnification.

To further investigate the impact of VPA treatment on cell pluripotency, we decided to measure the transcription activity of five genes implicated in pluripotency maintenance: *Oct4*, *Nanog*, *Klf4*, *Sox2* and *c-Myc* (**Figure 3.5**). Expression levels were initially analysed after 1mM VPA treatment for 8 hours, and then after 1 day or 5 days post wash out. For each time point, control untreated cells were harvested as well. Expression profiles revealed that after 8 hours of treatment the expression of *Klf4* and *Nanog* was significantly decreased by VPA while *Sox2* was significantly up-regulated. Transcription of *c-Myc* and *Oct4* was unaffected by the treatment, although there was a small decrease in *Oct4* transcript.

Very little effect was observed after 1 day wash-out. Moreover, the differences between controls and wash out samples were not greater than that observed between the various controls. Transcription levels in wash out samples could simply be the reflection of natural variations between asynchronous ES cell populations and we concluded that VPA treatment has no long-term effects on the expression of pluripotency-associated genes.

Lastly, the transcriptional activity of *c-Myc* in washed out cells was consistently maintained at the control level. However, the expression of this gene decreased substantially in wash out and control cultures compared to day 0. The transcription factor *c-Myc* is cell cycle-regulated which could explain this observation. c-Myc is expressed when the cell is actively dividing, promoting G1/S transition (Amati et al., 1998). As, in contrast to day 0 samples, when cells were harvested at 1 and 5 days after wash out they were confluent enough to be split. Similarly to what was observed during cell cycle analysis, a change in cell division activity due to normal culture condition variations could explain this change in *c-Myc* expression **(Figure 3.5)**.

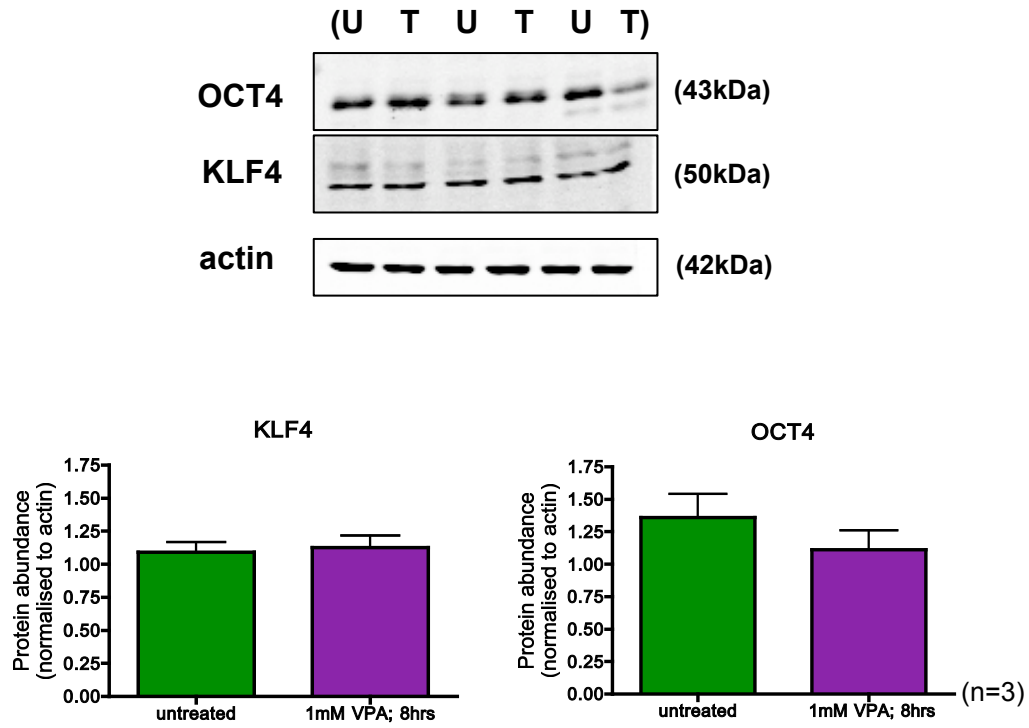
Figure 3.5 Effect of VPA treatment on transcription levels of pluripotency-associated genes



CCE/Rs were treated for 8 hours with 1mM VPA and then washed out and cultured for 1 or 5 days. RNA was extracted and quantified by RT-qPCR. The expression of five pluripotency associated genes was normalized against the housekeeper β-actin. A parametric two-tailed unpaired t-test showed significant differences for *Nanog*, *Klf4* and *Sox2* on the day of treatment (day0).

A drop in gene transcription doesn't always lead to a decrease in protein abundance. Therefore, we decided to use western blotting to quantify OCT4, NANOG and KLF4 in untreated cells and cells treated for 8 hours with 1mM VPA. Although *Klf4* transcription was decreased after the treatment, the amount of KLF4 protein remained unchanged following exposure to VPA. A very small decrease in *Oct4* transcript abundance was detected after the treatment and interestingly a slight effect was observed at a protein level, but this drop was non-significant. *Sox2* transcript abundance was increased following VPA treatment. Interestingly, the SOX2 protein has been found to be a cofactor of OCT4 (Avilion et al., 2003). Quantification of SOX2 protein level would be therefore relevant. Finally, NANOG protein abundance couldn't be assessed due to lack of availability of a specific antibody **(Figure 3.6)**.

Figure 3.6 Effect of VPA treatment on OCT4 and KLF4 protein abundance



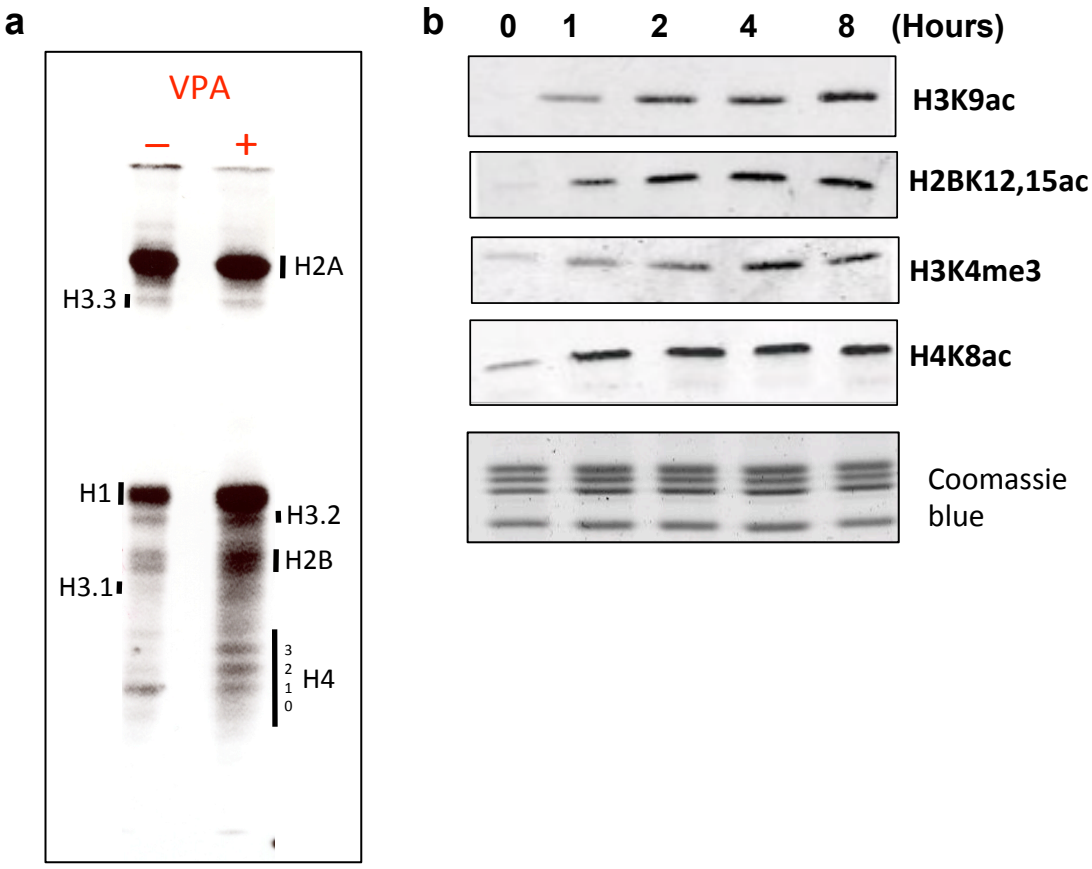
CCE/Rs were treated for 8 hours with 1mM VPA. Total protein was extracted and levels of OCT4 and KLF4 assessed using western blotting. β-actin was used for normalisation.
(U = untreated; T = treated)

3.1.3 Effect of valproic acid treatment on global histone modifications in ES cells

Once a non-toxic concentration of VPA has been established, we assessed the effect of the drug on global histone modifications. Initially we looked at global changes in acetylation level on specific histone isoforms. Histones extracted from untreated cells or cells treated for 8 hours with 1mM VPA were separated on a triton-acid-urea (TAU) gel and silver stained, allowing the visualisation of acetyl histone H4-isoforms. Histone H4 from untreated cells were predominantly mono-acetylated correlating with observations that chromatin in ES cells is hyperacetylated compared with chromatin in differentiated cells (Spivakov and Fisher, 2007). As expected, histone H4 from treated cells was found to be mostly di-, tri-acetylated **(Figure 3.7.a)**.

Using western blotting, specific histone modifications were then examined on histones extracted from cells that had been treated with 1mM VPA for 1, 2, 4 and 8 hours. As expected, the enrichment of each acetyl mark increased progressively during the time course, though to different extends, depending on the isoforms. Interestingly, methylation on H3K4 was increased as well in response to VPA treatment. This result confirmed a previously characterised cross talk between histone acetylation and H3K4me3, both marks being associated with active transcription (Nightingale et al., 2007) **(Figure 3.7.b)**.

Figure 3.7 Effect of VPA treatment on global histone modifications



a, Histones from untreated or treated (1mM; 8hrs) CCE/Rs were separated on a Triton-Acid-Urea (TAU) gel and then silver-stained. This type of gel allows the resolution of proteins according to mass and charge and thus, can resolve histone H4-acetyl isoforms (i.e. H4-1ac vs H4-2ac).
b, Histones were extracted from CCE/Rs that were treated with 1mM VPA for 1, 2, 4 and 8 hours. Changes in global levels of several histone modifications were assessed using western blotting. Coomassie blue staining was used as a loading control.

After we showed that 1mM VPA treatment induced general hyperacetylation of histones in our ES cell line, we investigated if those global changes could be heritable through mitosis. Because the turnover of histone modifications is very high, it was expected that if the cells couldn't retain hyperacetylation, abundance would drop very fast. Therefore, it was decided that cells would be washed out for short period of time at first. Cells were treated with 1mM VPA for 2, 4, and 8 hours. Following 8 hours of treatment some cells were washed out and cultured for a further 15 minutes, 30 minutes, 1, 2 or 4 hours. Histones were extracted at each time point and specific marks immunodetected by western blotting **(Figure 3.8.a)**. The secondary antibody was conjugated with a fluorescent dye allowing quantification of the signal. Each value was normalized using an antibody raised against the H3 C-terminal region **(Figure 3.8.b)**.

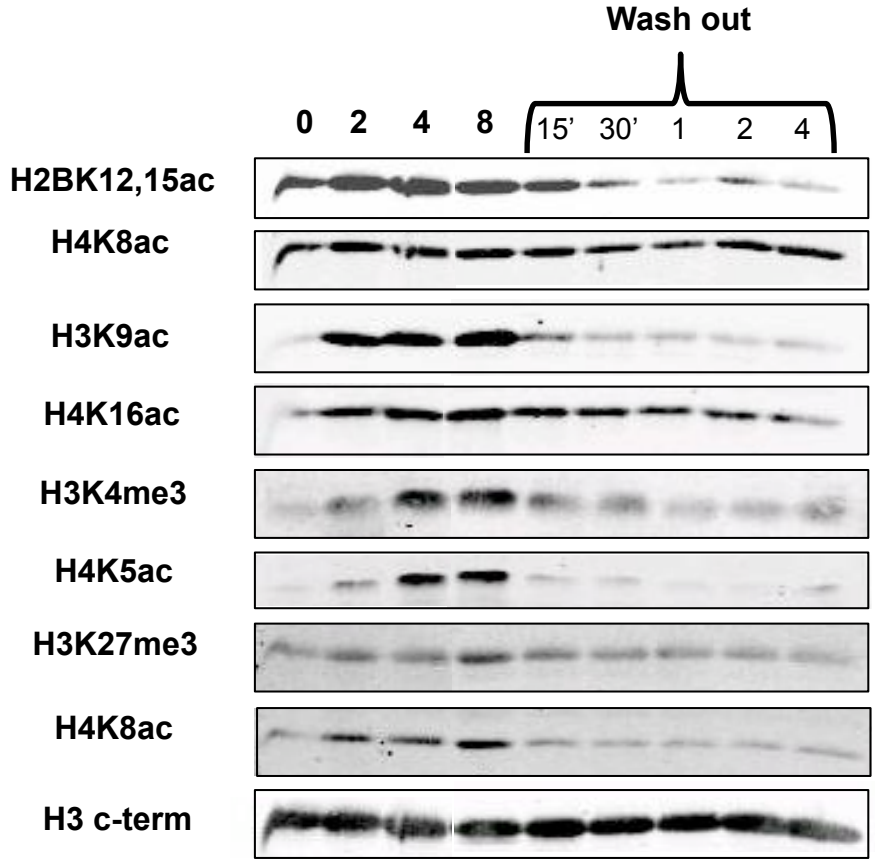
As expected, the abundance of each mark was increased by VPA treatment, reaching a plateau after 4 hours. The magnitude of the increase varied between the marks, for example the increase of H3K9ac (≈16 fold) was more significant than the increase of H2BK12,15ac (≈3 fold). The increase in acetylation on histones has previously been shown to be site specific in Hela cells, particularly on histone H3. With regard to histone H4, acetylation started at residue 16 and progresses in a N-terminal direction (i.e. mono-acetylation of H4 occurs on lysine 16 and further addition of acetyl groups occurs on lysine 12, 8 and 5) (Thorne et al., 1990; Turner et al., 1989). As a result, because of the shift to higher levels of acetylation following VPA treatment, the fold change of H4K16ac should be less substantial than the fold change of H4K5ac. Consistent with this, VPA treatment induced a (≈)15-fold increase in H4K5ac, a (≈)5-fold increase in H4K8ac and a (≈)3-fold increase in H4K16ac.

Finally, in contrast with H3K4me3, H3K27me3 abundance did not change in response to the treatment.

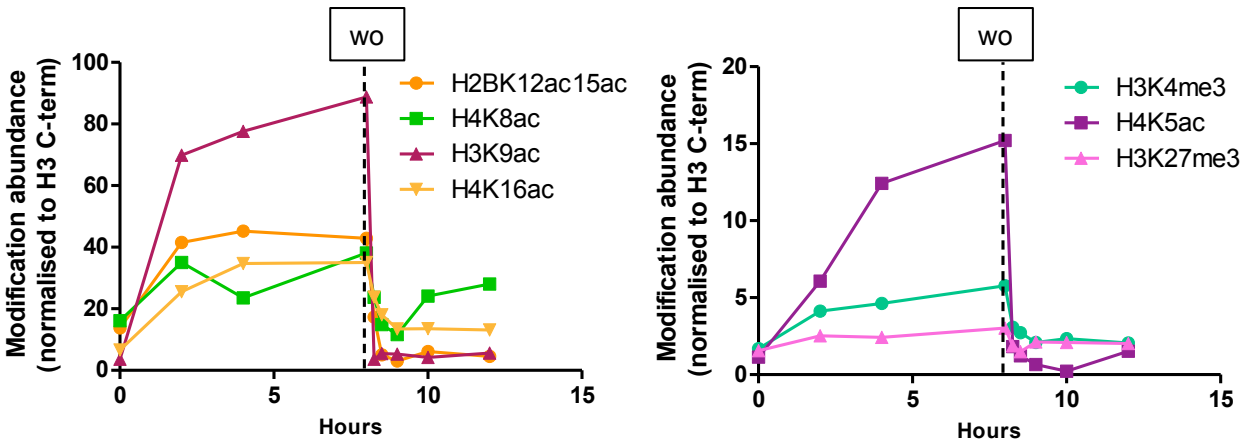
Removal of the drug after 8 hours of treatment induced a massive drop in histone acetylation and concomitantly of H3K4 methylation within only 15 minutes. This observation, consistent with a high histone acetylation turnover, very clearly showed that the cells were not able to retain the hyperacetylation, at least at a global level and for the modifications we looked at.

One possible explanation was that the cells hadn't been treated for long enough to keep the memory of hyperacetylation. Thus, we decided to repeat the experiment with a longer initial treatment. Cells were treated with 1mM VPA for 2, 4, 8, 12 and 16 hours before being washed out for 15 minutes, 30 minutes, 1, 2 or 4 hours. The same observations were made, with a loss of mark abundance within 15 minutes of wash out **(Figure 3.9)**.

Figure 3.8 Global histone modification abundance after VPA treatment followed by wash out

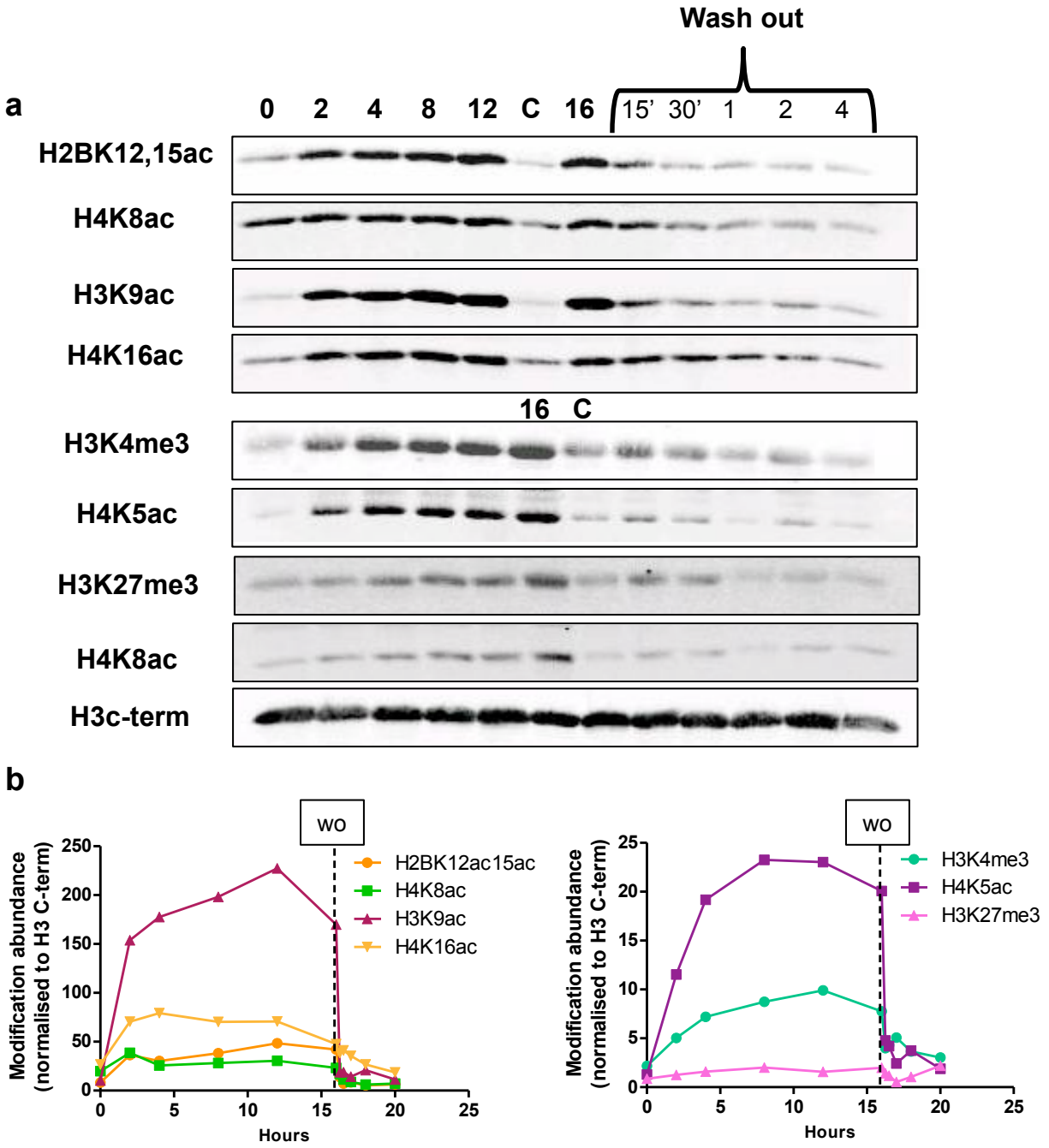


b



a, Histones were extracted from CCE/Rs that were treated with 1mM VPA for 0, 2, 4 and 8 hours followed by wash out (15mins, 30mins, 1,2 and 4 hours). Changes in global levels of several histone modifications were assessed using western blotting. Blotting with an H3 C-terminal antibody was used as a loading control.
b, Image analysis of **a**. H3 C-terminal staining was used to normalize values.

Figure 3.9 Global histone modification abundance after VPA treatment followed by wash out



a, Histones were extracted from CCE/Rs that were treated with 1mM VPA for 0, 2, 4, 8, 12 and 16 hours followed by wash out (15mins, 30mins, 1,2 and 4 hours). Changes in global levels of several histone modifications were assessed using western blotting. Blotting with an H3 C-terminal antibody was used as a loading control.

b, Image analysis of **a**. H3 C-terminal staining was used to normalize values.

3.1.4 Effect of valproic acid treatment on global gene expression in ES cells

VPA is an HDACi and as expected induced global hyperacetylation of histone in ES cells as well as an increase in H3K4me3. Those two modifications have been widely associated with active transcription so gene expression was likely to be affected after treatment. Microarray technology was used to compare global transcription activity before and after VPA treatment. In order to detect any persistent effect, gene expression in washed out cells was assessed as well.

Cells were treated for 8 hours with 1mM VPA and then harvested or washed out for 5 days. Two colour arrays were used to compare global gene expression in treated and control cells. cDNAs of treated and control cells were respectively labelled with the red Cy5 dye and the green Cy3 dye. M-values were obtained, representing the ratio between red and green fluorescent signals and therefore the differential expression in treated and control samples. The same experimental procedure was used to compare gene expression in wash out and control cells. For each condition, M-values from three independent experiments were tested for statistical significance using the web-software T-MeV and p values were corrected with the statistical environment "R" (Benjamini False Discovery Rate correction). After exposure to 1mM VPA for 8 hours, 1144 genes showed a significant change in transcription activity ($p \leq 0.1$ after FDR correction), including 674 up-regulated and 470 down-regulated genes **(Figure 3.10.a)**. However, after 5 days of wash out, transcriptional activity had reverted back to basal levels, as no change in expression between washed-out cells and control cells was statistically significant.

Twelve genes (6 up-regulated, 2 unchanged and 4 down-regulated) were quantified using qPCR in order to validate the microarray data. We decided to look at a

representative set of genes, more or less affected by the treatment. The heat map representing the triplicate data for each gene was generated, showing up-regulated genes in red and down-regulated genes in green. Fold changes were correlated with the brightness of the colour on the map **(Figure 3.10.b)**. These microarray data were confirmed by qPCR, showing different levels of up- and down-regulation after treatment, and no persistent effect after wash out **(Figure 3.11 and 3.12)**. We therefore concluded that VPA could not induce heritable changes of gene expression at least for the genes we examined in ES cells.

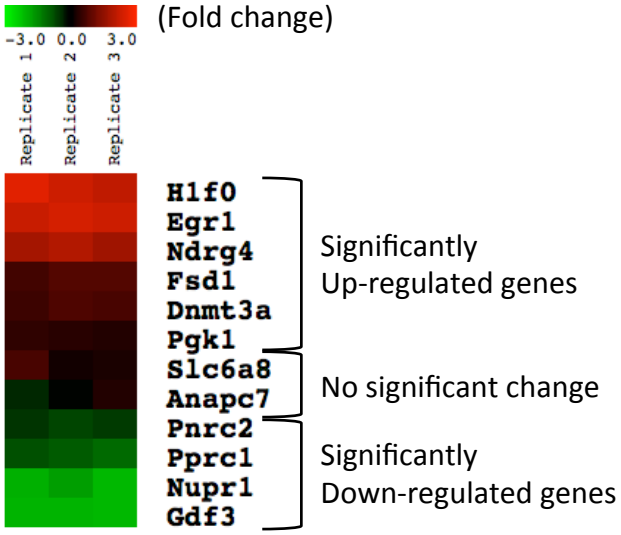
Figure 3.10 Effect of a 1mM VPA treatment on global gene expression

a

Change in expression	Up regulated	No change	Down regulated
Total	**674**	**9399**	**470**
≥ 2 fold	**95**		**28**
1.5-2 fold	**241**		**118**
1-1.5 fold	**338**		**324**

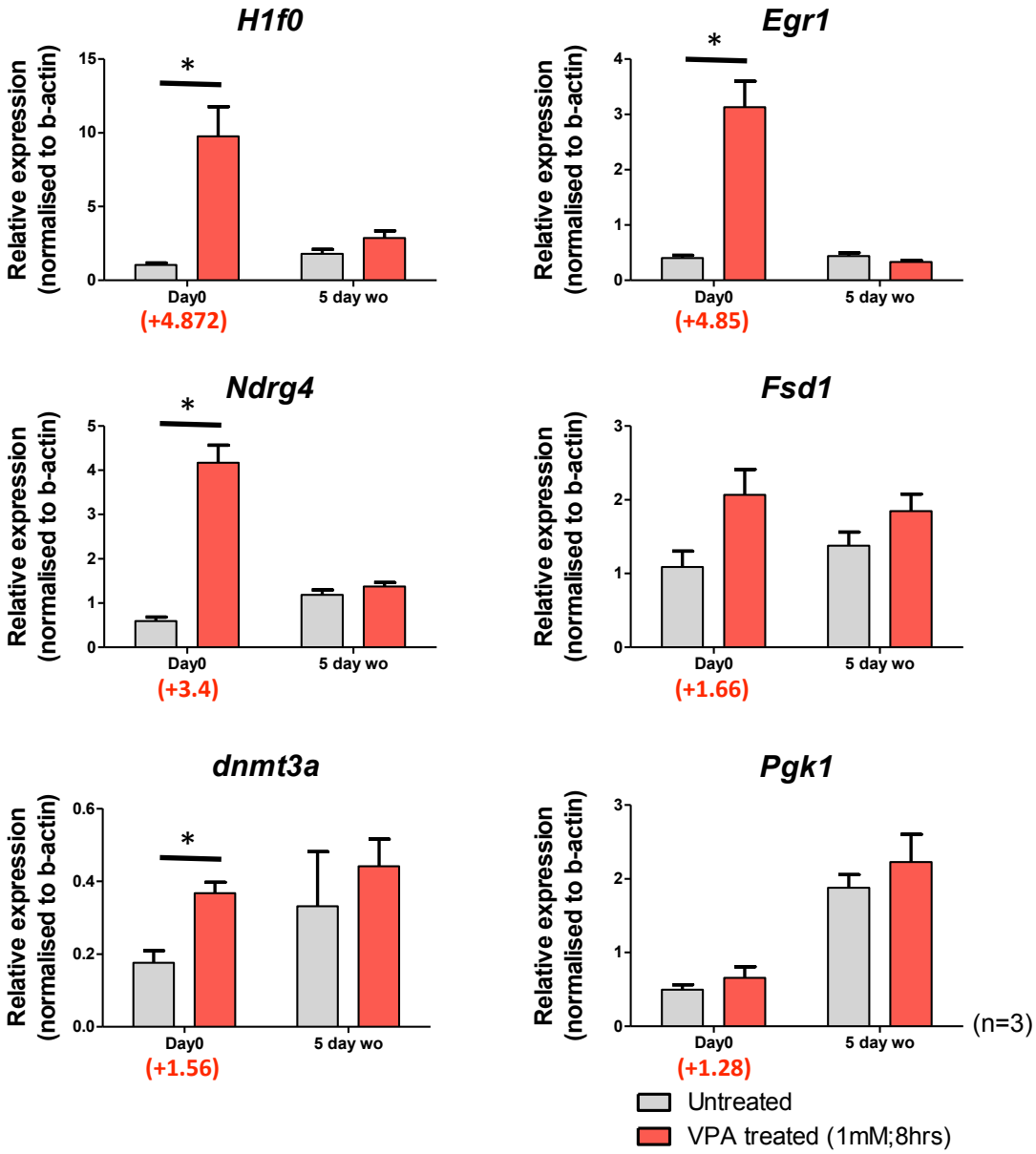
(n=3)

b



a, CCE/Rs were treated for 8 hours with 1mM VPA. RNA was extracted from control and treated cells and retrotranscribed. cDNA was then labelled using 2 different dyes for microarray hybridization. Values reflecting the ratio of fluorescence between the two dyes were obtained for each gene and normalized with the DNMAD web-software. Statistical tests were carried out with the web-software Tmev using three biological replicates and genes were classified as up-regulated, down-regulated or unchanged following VPA treatment. Genes were considered up- or down-regulated when the p value after FDR correction was equal or less than 0.1.
b, Heat map of the genes used for qPCR validation showing fold-changes between control and VPA treated cells. Red corresponds to a positive fold change in expression after treatment while green correspond to down-regulation of the gene. The larger the fold change the more bright the colour. In order to validate the microarray result we selected genes more or less affected in both groups as well as unchanged genes and tested them using qPCR.

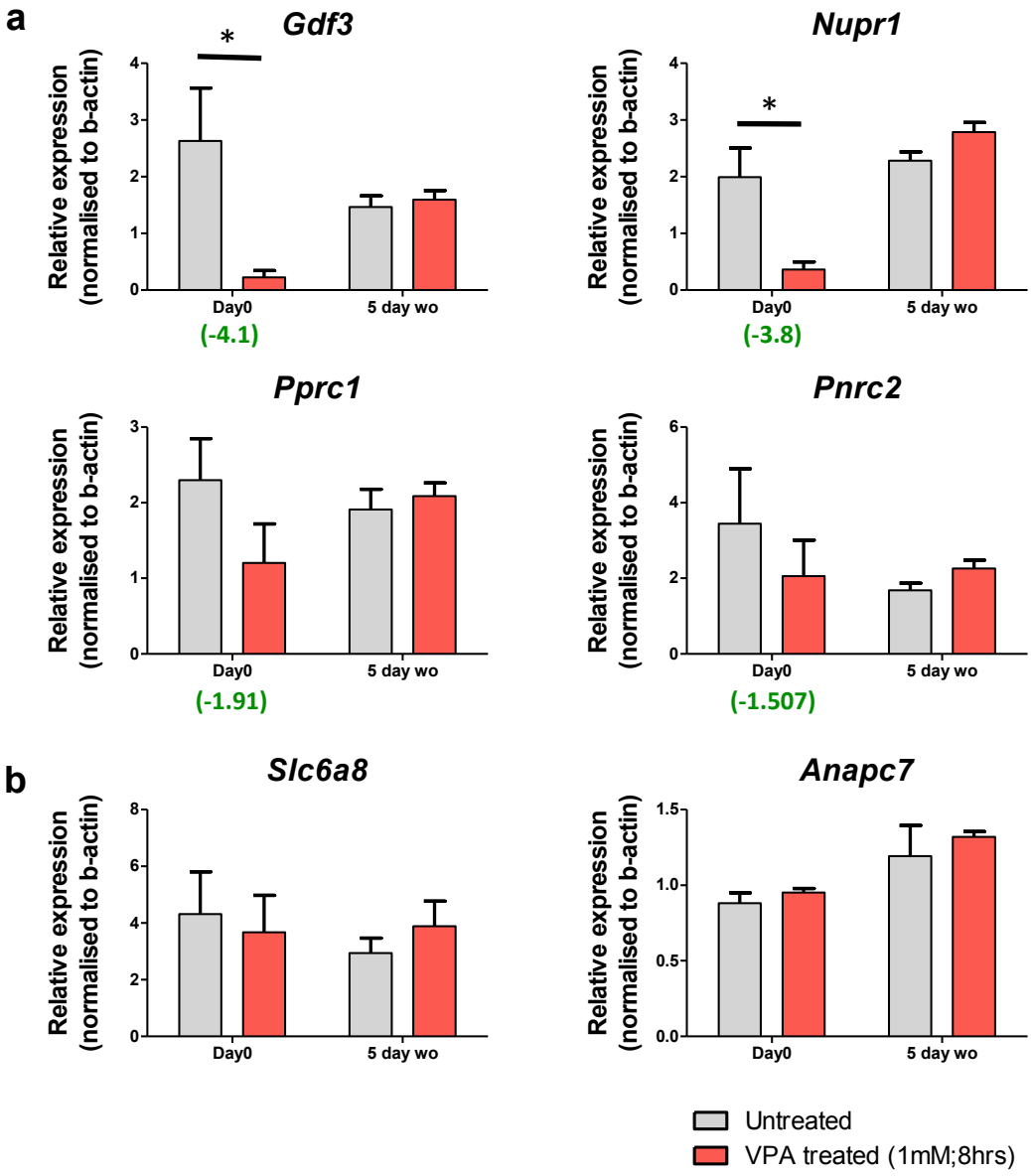
Figure 3.11 qPCR validation for up-regulated genes after VPA treatment



A group of genes significantly up-regulated after 8 hours of 1mM VPA treatment was selected from microarray data. Expression levels of each of those genes, after VPA treatment and after wash out, were obtained using qPCR for microarray validation. Values marked in red under each graph correspond to the average fold change obtained after microarray analysis.

Statistical analysis were processed with GraphPad software. Up-regulation of the genes in treated samples was tested using the nonparametric one-tailed Mann-Whitney test at 95% confidence interval. The same results were obtained when differences between untreated and treated samples was assessed with the parametric two-tailed unpaired t-test.

Figure 3.12 qPCR validation for down-regulated genes after VPA treatment



A group of genes significantly down-regulated (a) or un-affected (b) after 8 hours of 1mM VPA treatment were selected from microarray data. Expression levels of each of those genes, after VPA treatment and after wash out, were obtained using qPCR for microarray validation. Values marked in green under each graph correspond to the average fold change obtained after microarray analysis. Statistical analysis were processed with GraphPad software as before (see figure 3.11).

3.1.4.1 Gene cluster analysis

Because VPA is used for the treatment of diseases like cancer, it was interesting to explore the effect of the drug exposure on gene activity. This was even more relevant as ES cells were used for our study and the concept of cancer stem cells is now growing (Alison et al., 2011). The web software DAVID and the software Ingenuity™ were used for microarray data analysis. Those tools were used for the identification of gene clusters that were significantly enriched in a list of genes and thus, allowed the extraction of biological meaning from large data sets.

DAVID:

The 1144 genes affected by the VPA treatment were mapped to associated-biological annotations. The proportion of genes in each annotation from our list and from a background list (the entire mus musculus genome) were then compared. This comparison led to the identification of over-represented annotations in our data set **(Table 3.1.a)**. Interestingly, in the annotation with the highest enrichment score, 102 genes were connected with non-membrane-bound organelles including 57 genes participating in cytoskeleton organisation. This result was consistent with the morphological change of the cells observed after VPA treatment. Lysine acetylation is known to occur on important proteins for the formation and the movement of microfilaments like actin, cofilin, thymosin b10, profilin, moesin and tropomyosin. In addition, the treatment of epithelial cells with the HDACi TSA led to the accumulation of stable actin stress fibres (Kim et al., 2006). Our data have shown that cytoskeleton dynamics was also affected by VPA at a transcriptional level with the up-regulation of

β-actin, *Arpc2*, *Arpc5*, *tropomyosin*, *coronin*, *destrin*, *gelsolin*, *vinculin* and *laminA*. All of these genes are associated with cytoskeleton or cell mobility.

Although the treatment did not induce cell death, cell cycle arrest or differentiation, some associated genes show a change in expression after 8 hours in the presence of 1mM VPA. The cell cycle annotation was found to be over-represented in our data set with for example the down-regulation of *Ccnd1* and *Ccnd2* (Cyclin D1 and D2) and the up-regulation of *Ccne1* (Cyclin E1). Intriguingly, phosphorylation of the Rb (Retinoblastoma) protein by cyclin D is required to activate the E2F family of transcription factors, which then promotes the expression of cyclin E and therefore the G1/S transition (Coqueret, 2002). However, Rb-E2F was found to form a repressor complex at promoters of target genes through the recruitment of HDACs (Frolov and Dyson, 2004). This could explain the up-regulation of cyclin E despite the down-regulation of cyclin D. *Anapc1* was up-regulated by VPA treatment. The product of this gene is the subunit 1 of the anaphase promoting complex (APC) which targets cell cycle regulatory proteins to the proteasome through its ubiquitin ligase activity (Jorgensen et al., 2001). Interestingly, TSA treatment has been shown to enhance APC protein assembly rather than up-regulate the subunits (Kimata et al., 2008). Our data showed that the VPA treatment also induces down-regulation of genes that promote cell cycle progression like *Hcfc1* (Host cell factor C1) or *Mybl2* (Myeloblastosis oncogene-like 2). Hcfc1 is a chromatin bound protein that associates in complexes with HDAC or HMT to regulate cell proliferation (Tyagi et al., 2007; Wysocka et al., 2003). Similarly, the *Mybl2* transcription factor, was shown to interact with a protein complex that promotes the G2/M transition (Knight et al., 2009).

Very few genes associated with pluripotency were affected by VPA treatment. For example, microarray data confirmed the down-regulation of *Nanog* in addition to the down-regulation of *Dppa2* (developmental pluripotency associated 2) and the up regulation of *Creg1* (Cellular repressor of E1A-stimulated genes 1). *Dppa2* encodes a protein that inhibits the differentiation of ES cells and promotes proliferation (Du et al., 2010) while the product of *Creg1* represses E1A-stimulated genes involved in differentiation processes (Sacher et al., 2005).

The apoptosis annotation was not found to be over-represented in our data set. However, some associated genes were affected by the treatment. For instance, the expression of *Dap* (Death-associated protein), a protein involved in the autophagic cell death pathway (Bialik and Kimchi, 2009), was increased by VPA. The *Clusterin* gene, which encodes a protein with roles in apoptosis and survival during tumourigenesis (Rizzi and Bettuzzi, 2010), was up-regulated by VPA treatment. We also observed the up-regulation of several heat-shock proteins, pro-survival chaperones implicated in the cellular stress-response (Pratt et al., 2010).

Hyperacetylation induced by VPA treatment may alter chromatin through the up- and down-regulation of several factors involved in nucleosome formation and histone modification. The expression of several histone isoforms was modulated by the treatment. Expression of *H1f0* (H1 histone family member 0) and *H3f3b* (H3 histone, family 3B) was increased, while expression of *H2afy2* (H2A histone family, member Y2) was decreased. In addition, four different lysine-specific demethylases (*Kdm2a*, *Kdm3a*, *Kdm3b* and *Kdm6a*) were down-regulated by VPA while *Kdm1b* was up-regulated. *Myst2* (Histone acetyl transferase 2) was down regulated by VPA as well

as the lysine acetyltransferase *Kat2b*. Finally, DNA methylation was also potentially affected by the treatment as *dnmt3a* (DNA methyl-transferase 3a) was up-regulated.

The synthesis of a functional protein is a long process that requires multiple steps. Our study has shown that a VPA treatment induced changes in the transcriptional activity of genes involved in histone modifications, one of the major events that act on gene expression. In addition, we identified genes that participate in RNA processing and protein processing. *Cftf2t*, a germ cell expressed gene coding for a protein involved in polyadenylation of pre-mRNAs (Monarez et al., 2007), was found to be up-regulated. Transcription of *Gle1l*, a gene coding for a protein that participates in RNA transport to the cytoplasm (Stewart, 2007), was also increased after treatment. In addition, RNA processing was potentially altered by the down-regulation of *Hnrph2*, *Pabpn1*, *Tcerg1* and a series of splicing factors. HNRPH2, a heterogeneous nuclear ribonucleoprotein isoform, plays a suggested role in pre-mRNA processing, transport and interaction with nuclear structure (Honore et al., 1995). PABPN1 (Poly(A) binding protein nuclear 1) is part of a protein complex that ensures the elongation of poly(A) tails by poly(A) polymerase (Kuhn and Wahle, 2004). Finally, TCERG1 is a protein part of the spliceosome and interacts with RNA polymerase II to block its action (Banman et al., 2010).

Translation is the mechanism by which proteins are synthesised from mRNA. A group of five translation initiation factors and two translation elongation factors, small subunits of the ribosome (Rodnina and Wintermeyer, 2009), were found down-regulated after treatment. We also observed down regulation of *Iars*, a gene that encodes the isoleucyl-tRNA synthetase, an enzyme responsible for the attachment of

the amino acid isoleucyl to the corresponding tRNA (Shiba et al., 1994). On the other hand, *Rars* (Argininyl-tRNA synthetase) was up-regulated after treatment. Furthermore, VPA treatment induced a decrease in transcription of 5 ribosomal proteins. Translation in the mitochondria was also potentially affected by the treatment as *Mrps18b* (28S mitochondrial ribosomal protein S18B) was down regulated. Finally, expression of *Pdia6* (protein disulfide isomerase family A, member 6) coding for an enzyme that catalyses the formation of disulphide bonds in the endoplasmic reticulum (Jessop et al., 2009) was increased by VPA.

Table 3.1 Gene cluster analysis (DAVID; Ingenuity)

a

Cluster	E.S.	Examples of affected genes
Cytoskeleton/ cytoskeleton interracting proteins	4.29	*Actinβ, Arpc2, Coronin, Destrin, tropomyosin, coronin, destrin, gelsolin, vinculin, laminA.*
Transport	3.35	*Cldn6, Eea1, Dstn, Tcte1l, Ipo11, Kif3a, Kpnb1, Nsf, Sl4a2, Vdac1.*
Ubiquitine conjugation	3.31	*Nedd4, Psmb4, Rnf128, Mir, CathepsinL1, Smurf2.*
Cell cycle	3.17	*Ccnd1, Ccnd2, Ccne1, hcfc1, Mybl2, Anapc1.*
Vesicule formation	3.09	*Cltb, Ap2a1, Rab3d, Rad18, Scamp1, Itsn2.*
Lysosome metabolism	3.05	*Ctsl, Gm2a, Bcl10, Slc11a2, Manba.*
Transcription cofactors	2.41	*Nlk, Tbl1rx1, Trim28, Hcfc1, Ncoa1, Creg1, Maged1.*
Steroid/lipid metabolism	2.25	*Dpm1, Mvd, Pcyt1a, Lss, Agpat3.*
Chromatin organisation	2.19	*Dnmt3a, H1f0, H2afy2, H3f3b, Hira, Kdm1b-2a-3a-3b-6a, L3mbtl3, Myst2.*

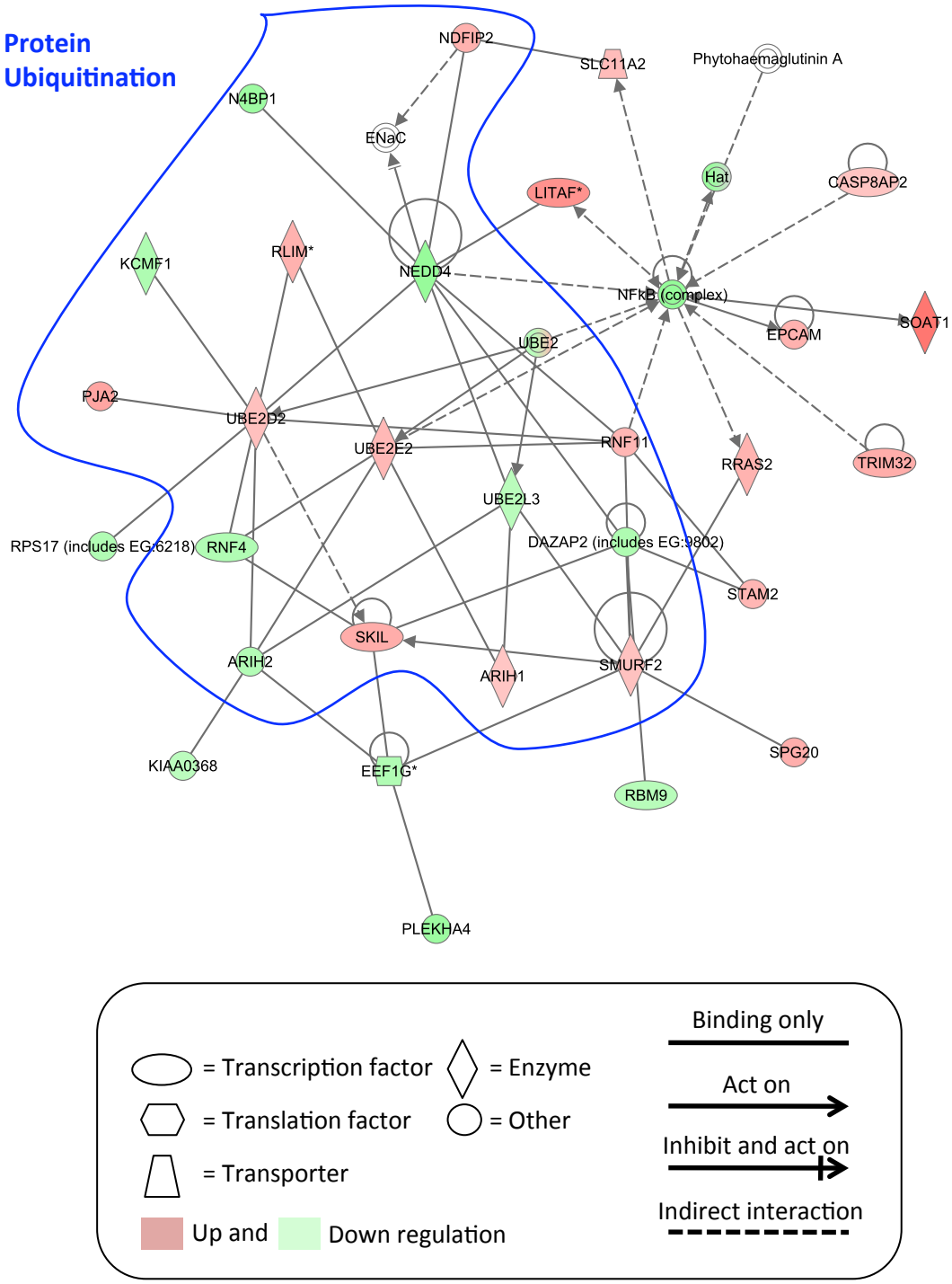
b

Network	Score
Cellular movement, assembly and organisation	50
RNA, Post-transcriptional modifications, cellular growth and proliferation	47
Protein degradation, synthesis and post-translational modifications	45
Development	44
Cell morphology	40

a, The bioinformatic software DAVID allowed the analysis of a large data set by mapping genes with associated biological annotations. Representation of each annotation, in our list of genes compared to the mouse genome, was given by enrichment scores. Each annotation contains several groups and the enrichment score is the geometric mean of all the group p-values (in minus log transformation). An enrichment score of 1.3 is equivalent to 0.05 in non-log scale, therefore any annotation with a enrichment equal or above 1.3 is significantly overrepresented in our data set. The table contains some of those relevant annotations.

b, The software Ingenuity was used for Network analysis. Genes were mapped together in a network when they were found to interact with each other in a biological context. For each network a score was generated, equivalent to the negative log of a p-value. A score of 2 indicated at 99% confidence that genes within a network did not associate due to random chance.

Figure 3.13 Example of a network built using Ingenuity



This network generated by Ingenuity was named " Protein degradation, synthesis and post-translational modifications". Red and green represent direction of change in our data set.

Ingenuity:

The software Ingenuity was subsequently used to build networks of interacting genes from our gene list. For each network a score was generated, equivalent to the negative log of a p-value. A score of 2 indicated at 99% confidence that genes within a network were not together due to random chance. Although the annotations were not as precise as the DAVID ones, Ingenuity highlighted similar functions, reinforcing confidence in the meaning of our data set **(Table 3.1.b)**. As an example, we chose to describe network 3, which showed interactions between proteins affected by VPA treatment and implicated in the ubiquitination pathway **(Figure 3.13)**. Ubiquitination of proteins is fundamental as it leads to degradation by the proteasome or lysosome but is also implicated in various processes like signalling or cell death. Three types of enzymes are required for the establishment of ubiquitin chains on substrates: ubiquitin-activating enzymes (E1s), ubiquitin-conjugating enzymes (E2s) and ubiquitin protein ligases (E3s) (Dikic et al., 2009). Many biological events require protein degradation, for instance cell cycle progression or the elimination of damaged proteins. NEDD4 (neural precursor cell-expressed developmentally downregulated gene 4) is an ubiquitin ligase that belongs to the HERC familly of E3s. This enzyme is implicated among others in the down-regulation of growth factors (Yang and Kumar, 2010) and the quantity of *Nedd4* transcript was decreased in VPA treated cells. However, transcription of *Ndfip2* (Nedd4 family interacting protein 2) was increased by the treatment. The corresponding protein enhances the activity of Nedd4 family-E3s and brings them to endosomes (Mund and Pelham, 2010). RLIM (Ring finger protein LIM domain interacting) is an E3 ubiquitine ligase that was identified as an important activator of X-chromosome inactivation. In fact it was shown to enhance

the expression of *Xist* in a dose-dependant manner. Interestingly, it was suggested that NANOG inhibits the transcription of *Rlim* (Barakat et al., 2010). This correlates with our data set where *Nanog* was down-regulated and *Rlim* up-regulated. VPA was previously shown to trigger HDAC2 degradation by the proteasome through up regulation of the E2 conjugating enzyme UBC8. However, RLIM protein level was found unaffected by VPA treatment (Kramer et al., 2003). The three additional ring-finger proteins *Rnf4*, *Rnf11* and *Pja2* were affected by VPA treatment. The gene *Rnf4*, down-regulated by VPA, encodes a Ring-finger protein that transfers ubiquitin on SUMOylated proteins. RNF4 acts on genome stability and enhances DNA demethylation (Hu et al., 2010). *Rnf11* was on the other hand up-regulated by the treatment. This gene is up-regulated in some tumours and was demonstrated to be essential for the activity of A20, a negative regulator of NF-kB (Shembade et al., 2009). E3s ubiquitin-ligase enzymes interact through the HECT domain with E2s ubiquitin-conjugating enzymes like UBE2D2 or UBE2E2 (Kamadurai et al., 2009) that were up-regulated by VPA treatment. As mentioned earlier, ubiquitination is implicated in a multitude of processes and therefore ubiquitin enzymes have a lot of substrates. Some of them were affected by the treatment as seen in the network. For instance, the RING finger ubiquitin ligase ARIH2 inhibits cell proliferation for myeloid differentiation (Marteijn et al., 2009) and was down-regulated in VPA-treated cells. *Skil*, up-regulated by VPA, encodes a negative regulator of TGFβ signalling. For example, SKIL blocks TGFβ-induced expression of ADAM12, a gene often found up-regulated in cancer or cardiovascular diseases. Upon TGFβ stimulation, SKIL is ubiquitinated by SMURF2 and degraded by the proteasome (Solomon et al., 2010).

We could therefore imagine here that the up-regulation of *Skil* was balanced by the up-regulation of *Smurf2*.

We observed that VPA treatment affected about 10% of genes and those genes were associated with many biological functions like cytoskeleton dynamics, ubiquitin conjugation or cell cycle progression. However, fold changes were not very large and a change in transcriptional activity does not always affect protein abundance. For instance, our FACS experiments have previously shown that the cell cycle of VPA treated cells was not extensively disrupted. Therefore we concluded that an 8 hour exposure to 1mM VPA induced consistent but small effects on gene expression in ES cells. In addition, the changes observed after treatment were not persistent when the drug was removed from the media.

3.1.5 Effect of valproic acid treatment on local histone modifications in ES cells

Although VPA induced fast and substantial hyperacetylation of histones in ES cells, only a small proportion of genes showed a change in their transcriptional activity. We therefore used chromatin immunoprecipitation to examine local histone modifications at specific loci before and after treatment with 1mM VPA for 8 hours. Chromatin was extracted from cells and precipitated with different antibodies raised against four specific histone marks: three marks associated with active transcription H3K9ac, H4K8ac, H3K4me3 and the silencing mark H3K27me3. DNA from bound and unbound fractions was analysed using qPCR. Sets of primers were designed in order to look at specific regions in the promoters of several genes. A ratio of B/UB greater than 1 indicates that the mark is enriched at that locus. We wanted to investigate whether the hyperacetylation of histones occurred uniformly throughout the genome or only at genes showing transcriptional change. We decided to look at the three up-regulated genes *Egr1*, *Ndrg4* and *H1f0* **(Figure 3.14-16)**, two unaffected genes *Anapc7* and *Slc6a8* **(Figure 3.17-18)** as well as the down-regulated gene *Nanog* **(Figure 3.19)**. As expected in untreated cells, the enrichment of the different marks was for most cases found in the promoter (i.e. Egr1/2 and 3).

The first important observation concerned acetylation on H3K9 that appeared to increase at the promoter of all the genes we looked at. Interestingly, levels of enrichment in untreated cells didn't correlate with fold changes after treatment. Importantly, hyperacetylation of H3K9 at those loci were not related to the transcriptional change induced by VPA, and the fold changes on individual promoters were not as high as expected. In fact, when bulk histones were analysed by western blot we observed a 16-fold increase in H3K9ac. But, surprisingly, at specific regions

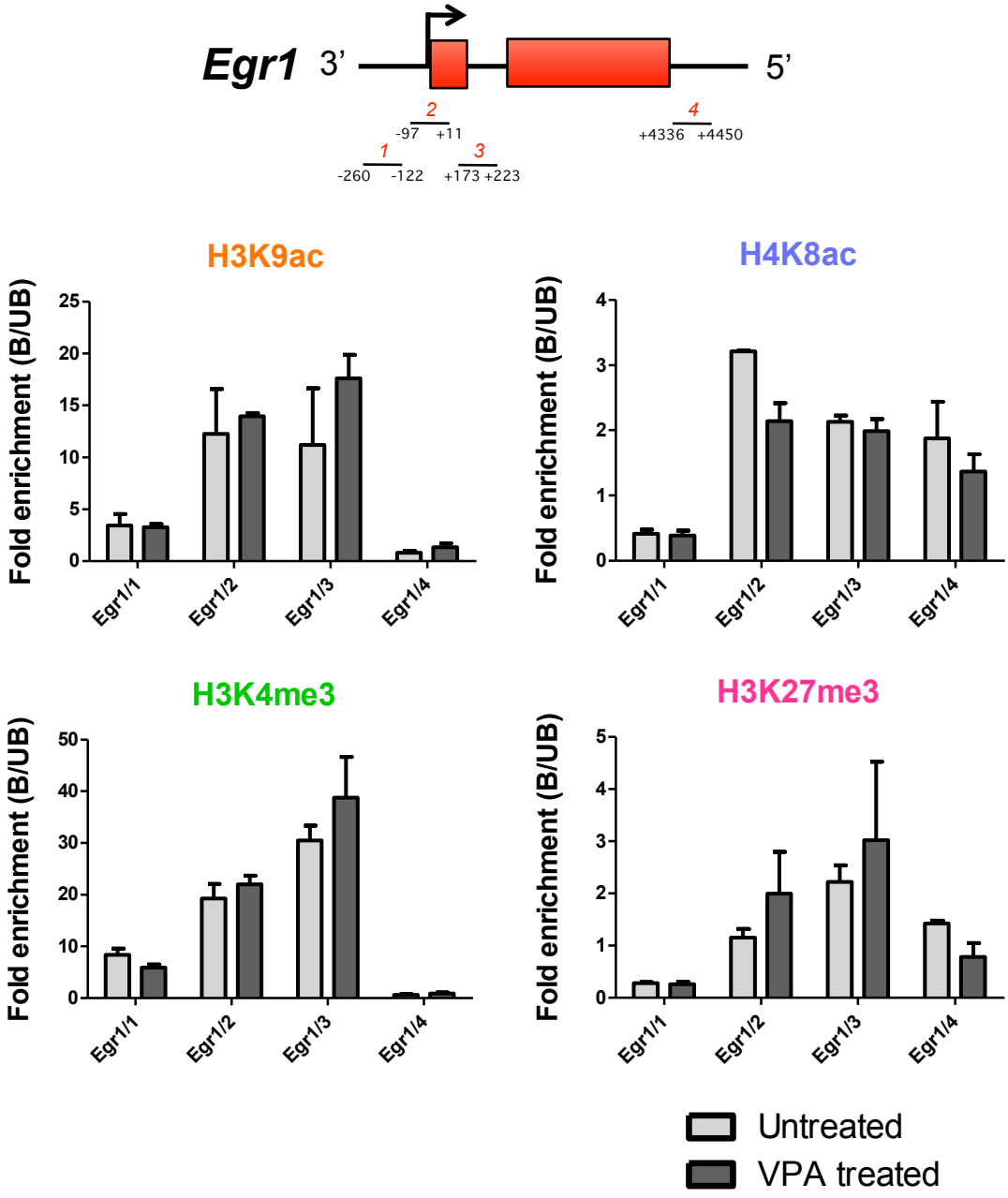
of the promoters we studied by ChIP, the increase in H3K9ac was never greater than 2-fold.

While H3K9ac generally increased after treatment, acetylation on H4K8 consistently decreased over the promoters of the genes we looked at. Surprisingly, but consistent with this result, a recent study showed that VPA treatment induces significant deacetylation of H3 and H4 (Rada-Iglesias et al., 2007).

As expected, H3K4me3 was affected by VPA treatment but not uniformly. Decrease in enrichment after treatment was observed at the promoter of *Nanog* and *Anapc7* and very modestly at the promoter of *Slc26a8*. Interestingly, increase in H3K4me3 was only observed at the up-regulated genes *Egr1* and *Ndrg4*. However, no change was observed for the *H1f0* gene. H3K4me3 could play a major role in the VPA response but more genes should be analysed for final conclusions.

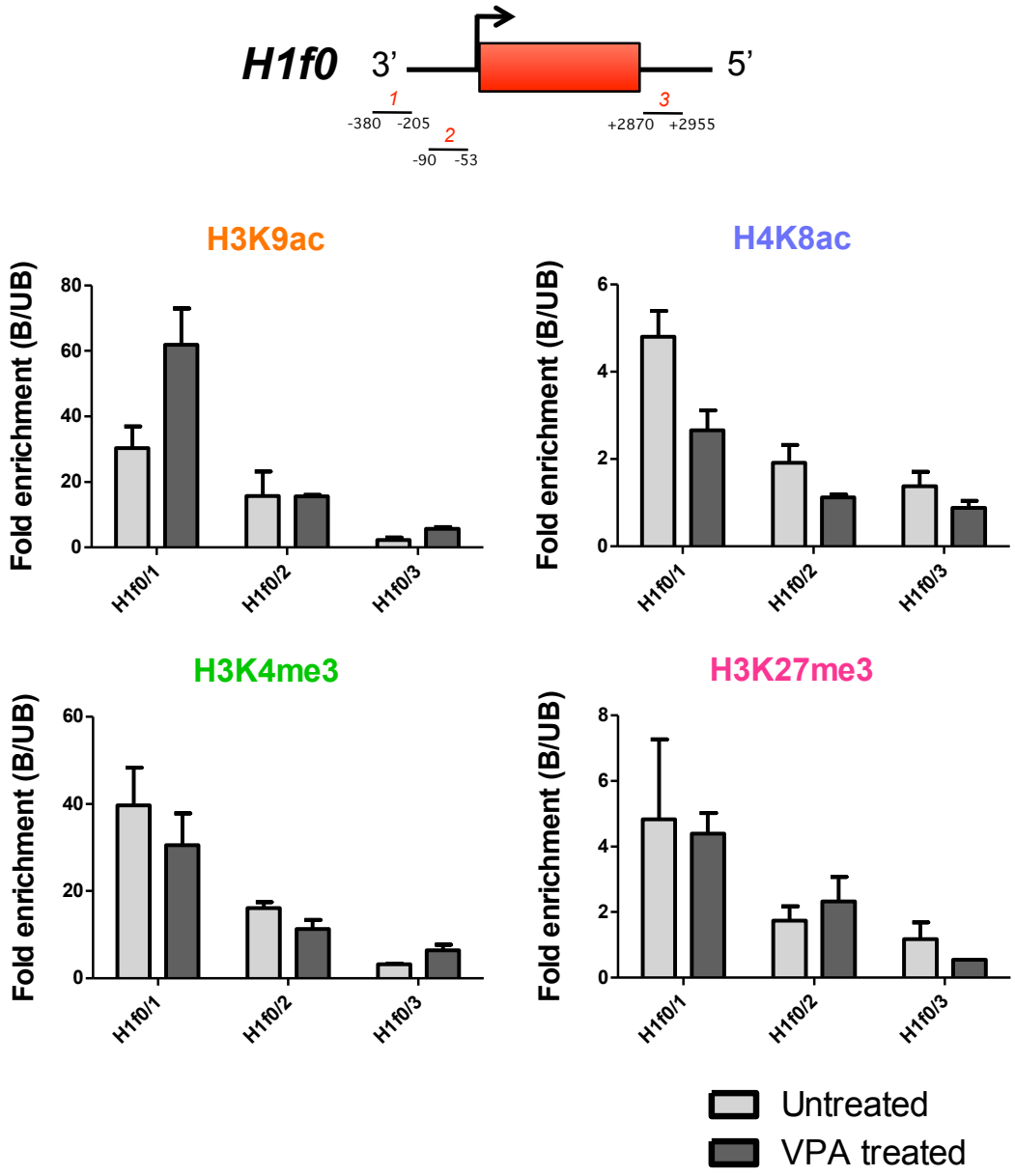
Finally, consistently with the data obtained by western blotting, most of the genes we looked at were not enriched for H3K27me3 and this feature was unaffected by the treatment. The mark was present on the promoters of *Egr1* and *H1f0* only and slightly increased after VPA treatment on *Egr1*, although H3K27me3 is a silencing mark and *Egr1* was up-regulated by the treatment.

Figure 3.14 Distribution of histone modifications on the *Egr1* gene before and after VPA treatment



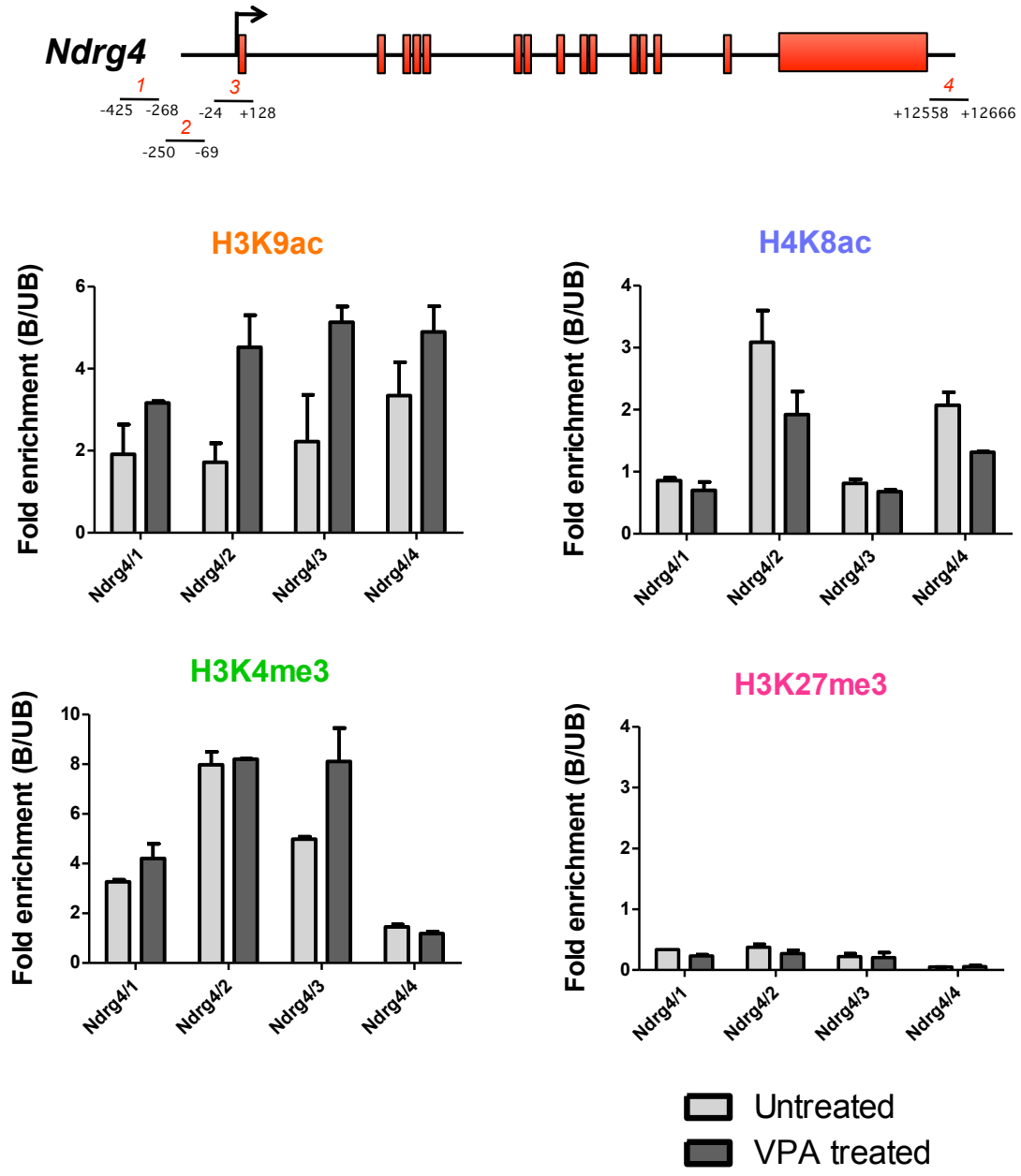
CCE/Rs were treated with 1mM VPA for 8 hours. Chromatin was extracted from control and treated cells and immunoprecipitated with antibodies raised against three histone modifications associated with active transcription (H3K9ac, H4K8ac, H3K4me3) and one associated with repressed transcription (H3K27me3). DNA in bound (B) and unbound (UB) fractions were analysed by qPCR using region-specific primers. The relative enrichment was calculated as the ratio B/UB and the mark was considered as present if this ratio was equal to or greater than 1. Error bars represent the standard error of the mean (n=2).

Figure 3.15 Distribution of histone modifications on the *H1f0* gene before and after VPA treatment



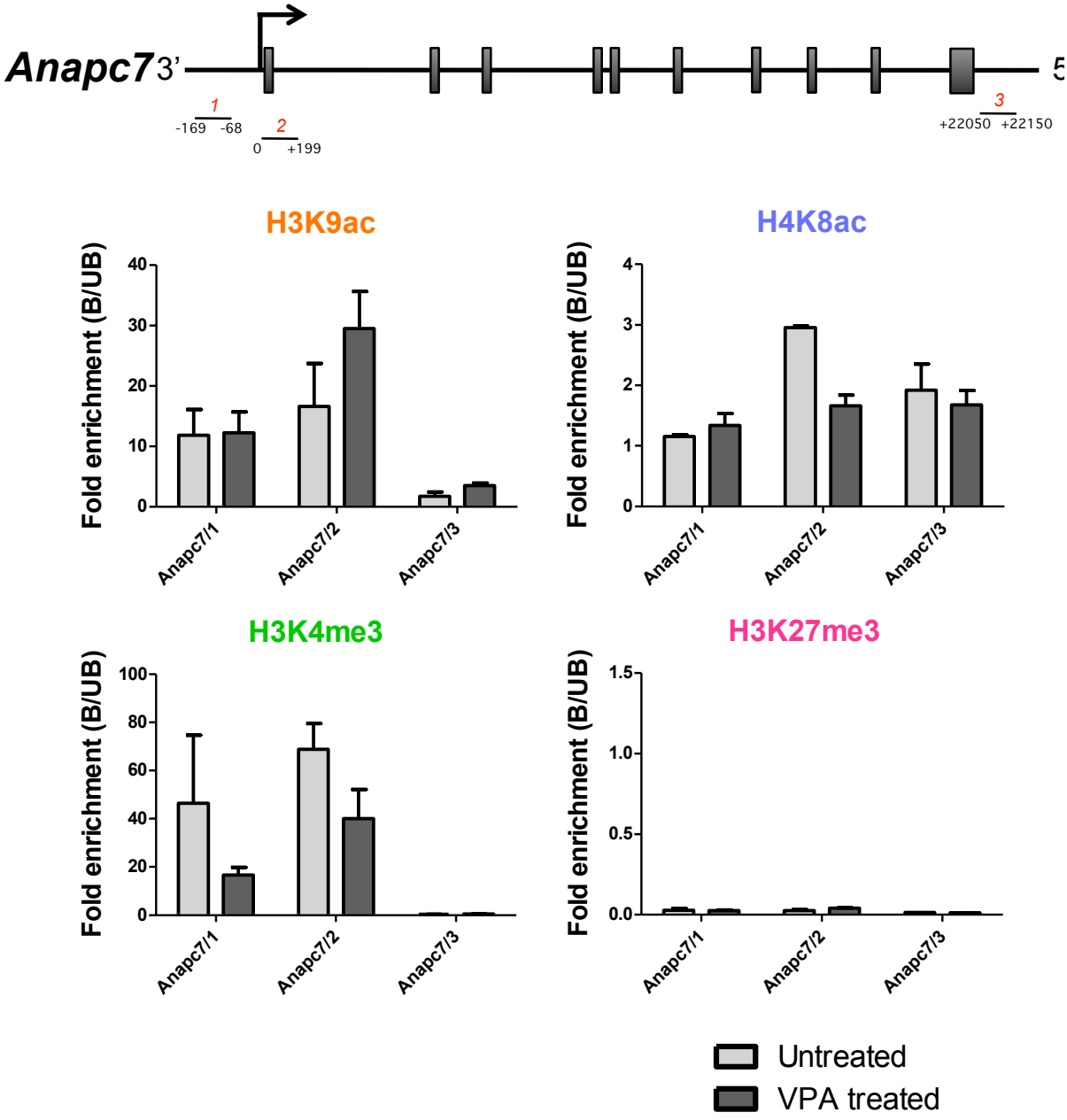
CCE/Rs were treated with 1mM VPA for 8 hours. Chromatin was extracted from control and treated cells and immunoprecipitated with antibodies raised against three histone modifications associated with active transcription (H3K9ac, H4K8ac, H3K4me3) and one associated with repressed transcription (H3K27me3). DNA in bound (B) and unbound (UB) fractions were analysed by qPCR using region-specific primers. The relative enrichment was calculated as the ratio B/UB and the mark was considered as present if this ratio was equal to or greater than 1. Error bars represent the standard error of the mean (n=2).

Figure 3.16 Distribution of histone modifications on the *Ndrg4* gene before and after VPA treatment



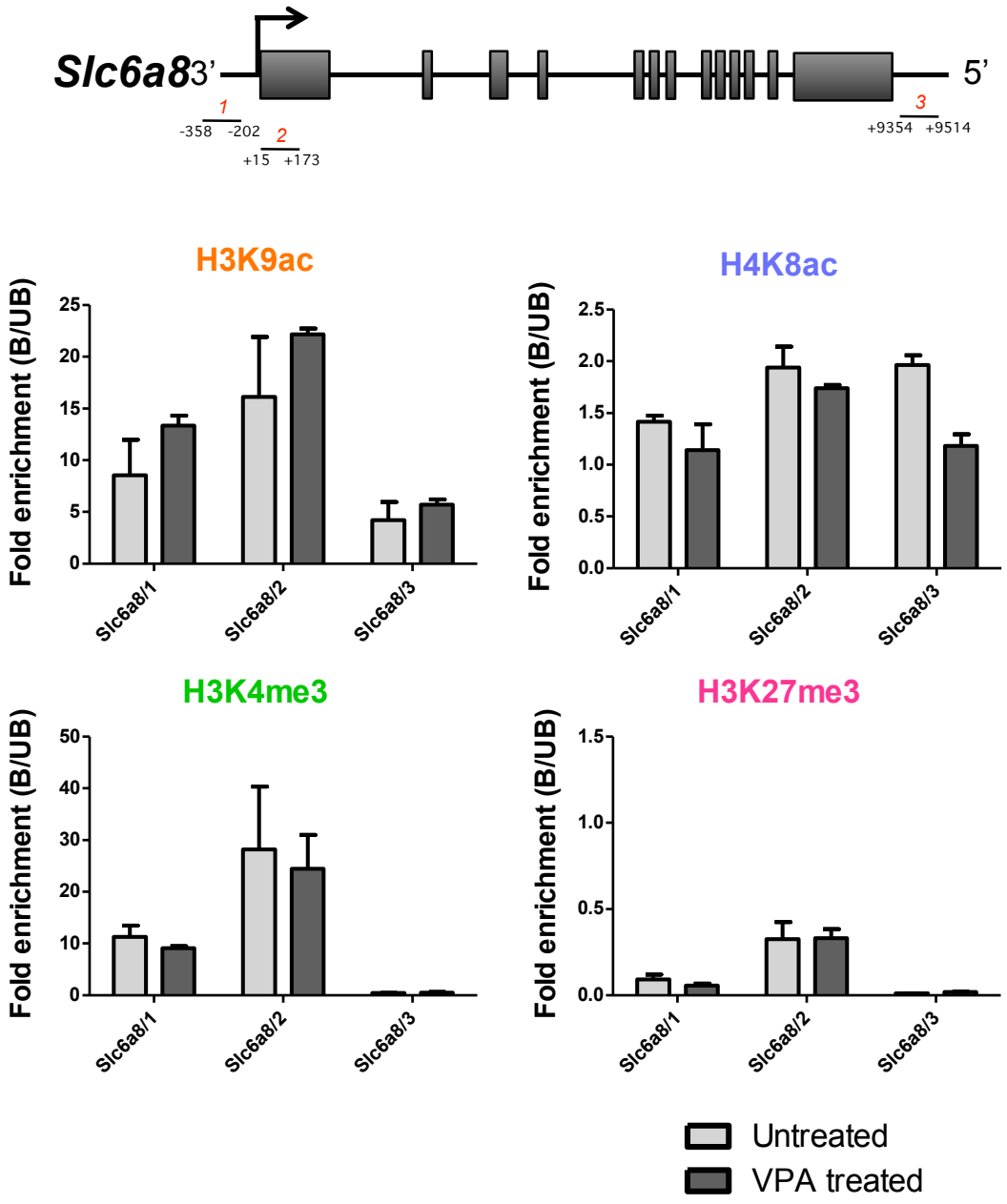
CCE/Rs were treated with 1mM VPA for 8 hours. Chromatin was extracted from control and treated cells and immunoprecipitated with antibodies raised against three histone modifications associated with active transcription (H3K9ac, H4K8ac, H3K4me3) and one associated with repressed transcription (H3K27me3). DNA in bound (B) and unbound (UB) fractions were analysed by qPCR using region-specific primers. The relative enrichment was calculated as the ratio B/UB and the mark was considered as present if this ratio was equal to or greater than 1. Error bars represent the standard error of the mean (n=2).

Figure 3.17 Distribution of histone modifications on the *Anapc7* gene before and after VPA treatment



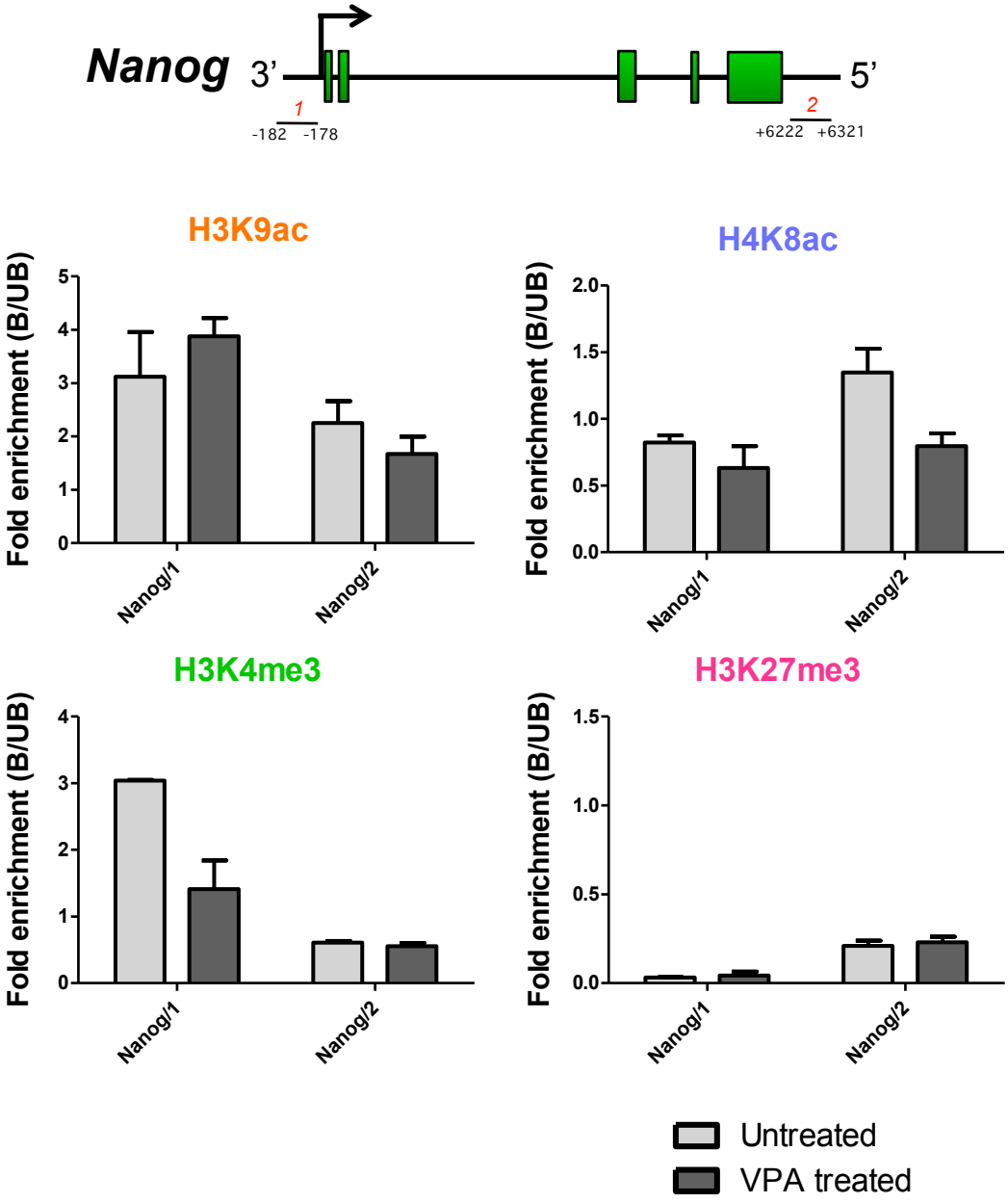
CCE/Rs were treated with 1mM VPA for 8 hours. Chromatin was extracted from control and treated cells and immunoprecipitated with antibodies raised against three histone modifications associated with active transcription (H3K9ac, H4K8ac, H3K4me3) and one associated with repressed transcription (H3K27me3). DNA in bound (B) and unbound (UB) fractions were analysed by qPCR using region-specific primers. The relative enrichment was calculated as the ratio B/UB and the mark was considered as present if this ratio was equal to or greater than 1. Error bars represent the standard error of the mean (n=2).

Figure 3.18 Distribution of histone modifications on the *Slc6a8* gene before and after VPA treatment



CCE/Rs were treated with 1mM VPA for 8 hours. Chromatin was extracted from control and treated cells and immunoprecipitated with antibodies raised against three histone modifications associated with active transcription (H3K9ac, H4K8ac, H3K4me3) and one associated with repressed transcription (H3K27me3). DNA in bound (B) and unbound (UB) fractions were analysed by qPCR using region-specific primers. The relative enrichment was calculated as the ratio B/UB and the mark was considered as present if this ratio was equal to or greater than 1. Error bars represent the standard error of the mean (n=2).

Figure 3.19 Distribution of histone modifications on the *Nanog* gene before and after VPA treatment



CCE/Rs were treated with 1mM VPA for 8 hours. Chromatin was extracted from control and treated cells and immunoprecipitated with antibodies raised against three histone modifications associated with active transcription (H3K9ac, H4K8ac, H3K4me3) and one associated with repressed transcription (H3K27me3). DNA in bound (B) and unbound (UB) fractions were analysed by qPCR using region-specific primers. The relative enrichment was calculated as the ratio B/UB and the mark was considered as present if this ratio was equal to or greater than 1. Error bars represent the standard error of the mean (n=2).

3.1.6 Summary

A model system was designed to study short and long-term effects of a VPA treatment in ES cells. Firstly we looked at the effect of a VPA treatment on the cell cycle profile, viability and pluripotency of ES cells. We decided to use VPA at a concentration of 1mM and showed that it didn't significantly affect the cell cycle or the viability of the cells. On the other hand, VPA appeared to have an effect on pluripotency, inducing a change in cell morphology and a decrease of *Oct4*, *Nanog* and *Klf4* transcripts. However, we could not observe any significant effect at a protein level for Oct4 and Klf4. Importantly, after 5 days of wash out, cells displayed a normal cell cycle profile, regained typical morphology and expressed pluripotency-associated genes.

Western blot analysis confirmed that 1mM VPA treatment for 8 hours induced global hyperacetylation of histones in ES cells. Hyperacetylation was shown to occur on all histone isoforms and in a site-specific manner. Immunoblotting of different histone marks reasserted the existence of cross-talk between two active marks, H3 acetylation and tri-methylation of H3K4. In fact, the increase in H3K4me3 was concomitant with the hyperacetylation induced by VPA treatment.

Short time wash out experiments revealed that within only 15 minutes after removal of VPA, cells recovered normal levels of acetylation. Increasing the time of treatment up to 16 hours did not alter this result, as rapid loss of acetyl-marks was still observed. At a global level, and at least for the marks we examined, we could conclude that cells were not able to retain hyperacetylation in the absence of VPA. However, as the western blot technique can not resolve modifications at individual genes or at a local level (e.g. on the promoter of a specific gene), it was not possible

to assess the distribution and/or inheritance of hyperacetylation after VPA removal. Microarray analysis revealed that the VPA treatment did not extensively affect transcription activity. Indeed, only 10% of the genes were affected with almost equal number showing up- and down-regulation. Fold changes were not substantial, however we could detect specific groups of genes affected by the treatment. VPA-induced changes were not retained post wash-out. We therefore used ChIP experiments to assess histone modification abundance at specific regions of defined genes before and after VPA treatment. Because we looked at a small number of genes and qPCR only allowed the analysis of small regions of the genome, it was difficult to draw general conclusions from those experiments. However, we couldn't detect any pattern of histone modifications induced by VPA treatment that was specific to up- or down-regulation. In fact, an increase in H3K9ac was always observed at promoters after VPA treatment. Furthermore, although western blot experiments revealed that H4K8ac abundance increased following exposure to VPA, down-regulation of the mark was observed on promoters of individual genes. Finally, fold changes observed by ChIP did not reflect what was observed using western blotting. This phenomenon may suggests the existence of mechanisms that protect genes against hyperacetylation.

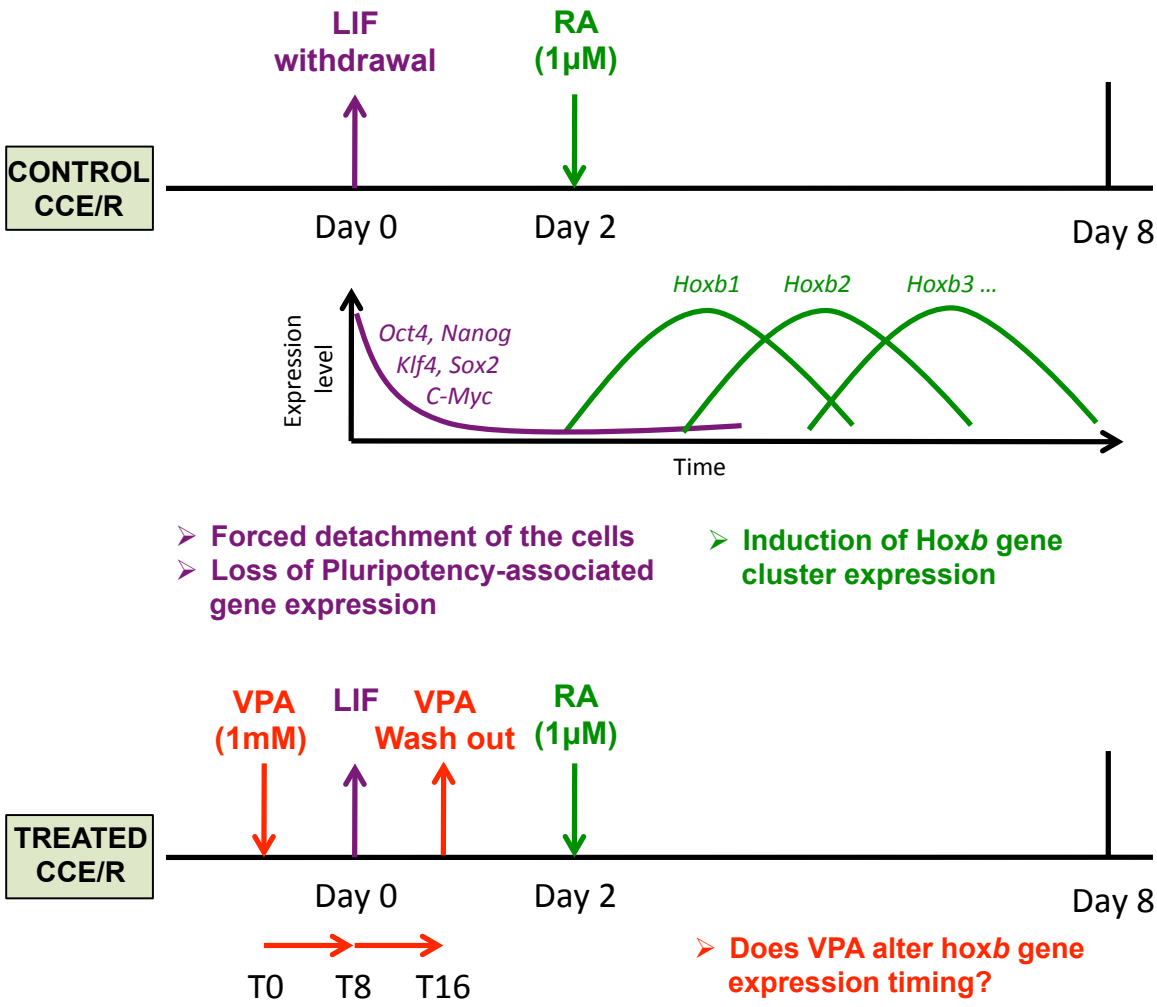
3.2 Effect of histone hyperacetylation on *hoxb* gene expression patterns during differentiation of ES cells

3.2.1 Introduction

Our first model system was intended to investigate the short and long term effect of histone hyperacetylation induced by the exposure to an HDACi, VPA. Measurement of gene transcription activity and histone modification levels showed that a time-restricted VPA treatment wasn't enough to induce heritable effects in ES cells. Although this result indicated that histone modifications on their own are not likely to be causative of long-term changes in genome function, it didn't prove that they couldn't be predictive of transcription. In fact, histone modifications are probably only one layer of all the many signals participating in the establishment of specific gene expression patterns. In order to test this hypothesis, we designed a second model system involving differentiation of ES cells and VPA treatment **(Figure 3.20)**. ES cells were replated in non-adherent dishes for the formation of embryonic bodies **(Figure 3.21)**. Differentiation was induced by the withdrawal of the pluripotency signalling factor LIF (Leukaemia Inhibitory Factor) and Retinoic Acid (RA) treatment, which was essential for the induction of *hoxb* gene expression. RA is a derivative of vitamin A, and plays an important role as a cell-signalling molecule during organogenesis, notably for the development of the hindbrain, spinal cord, heart, eye and other organs (Duester, 2008). RA acts as a ligand for the nuclear receptors RARα,β,γ and RXRα,β,γ which then bind to response elements present in the promoter of target genes (Rochette-Egly and Germain, 2009). Two RA response elements (RAREs) have been identified 3' of the *Hoxb1* gene, participating in the induction of gene expression in the *hoxb* cluster in response to retinoid exposure (Langston et al.,

1997). In normal conditions, RA induces *hoxb* gene expression in a very specific manner, in which the 10 *hoxb* genes (*hoxb1-9;13*) are successively expressed during differentiation following their physical order on the cluster. A protocol allowing ES cell differentiation in culture using RA has been optimized (Chambeyron and Bickmore, 2004). Based on this protocol, 1μM RA was added to the medium 2 days after LIF withdrawal and the differentiation of CCE/R cells was conducted for 8 days. In a first experiment cells were only treated with RA and samples were harvested every day for RNA extraction and quantification. In a second experiment, cells were differentiated using the same protocol with the addition of 1mM VPA to the culture medium for 16 hours (8 hours before and after LIF withdrawal). We hypothesised that histone hyperacetylation combined with the signal to differentiate could impact on *hoxb* gene expression timing. Thus, RNA was extracted every day during eight days and *hoxb* gene transcripts were quantified using qPCR **(Figure.3.20)**.

Figure 3.20 Model system #2:
Can VPA impact on the expression timing of *Hoxb* genes during ES cell differentiation?

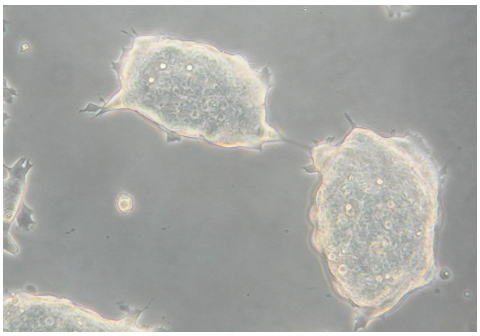


The model system was made to investigate the effect of hyperacetylation on the *hoxb* gene expression timing during differentiation. The *hoxb* gene cluster contains 10 genes that are successively induced in the developing embryo. Genes within the cluster are arranged in their temporal order of expression during differentiation as well as their expression domains along the antero-posterior axis. It was possible to reproduce the co-linear timing of *hoxb* gene expression in culture. For that, ES cells received 3 signals: on day 0 (1) detachment from the flask for the formation of embryonic bodies; (2) withdrawal of the pluripotency associated signal molecule LIF; and 2 days later (3) addition of 1µM retinoic acid in the medium.

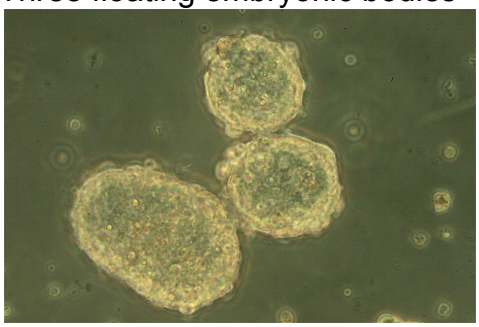
CCE/Rs were first differentiated using this protocol and several *hoxb* genes were quantified with qPCR. In a second experiment 1mM VPA was added into the medium for 16 hours overlapping with the first two signal of differentiation. Expression of hoxb genes was then quantified in order to measure the effect of the initial hyperacetylation.

Figure 3.21 Cell morphology during differentiation

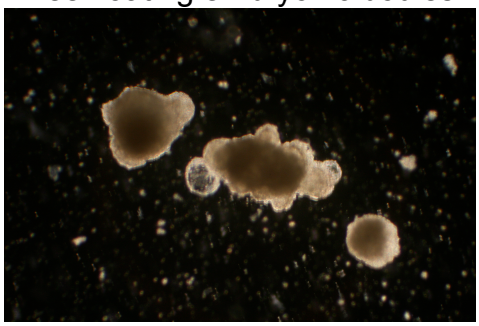
Undifferentiated CCER:
Two colonies attached to the flask



Differentiated CCER (Day1):
Three floating embryonic bodies



Differentiated CCER (Day7):
Three floating embryonic bodies



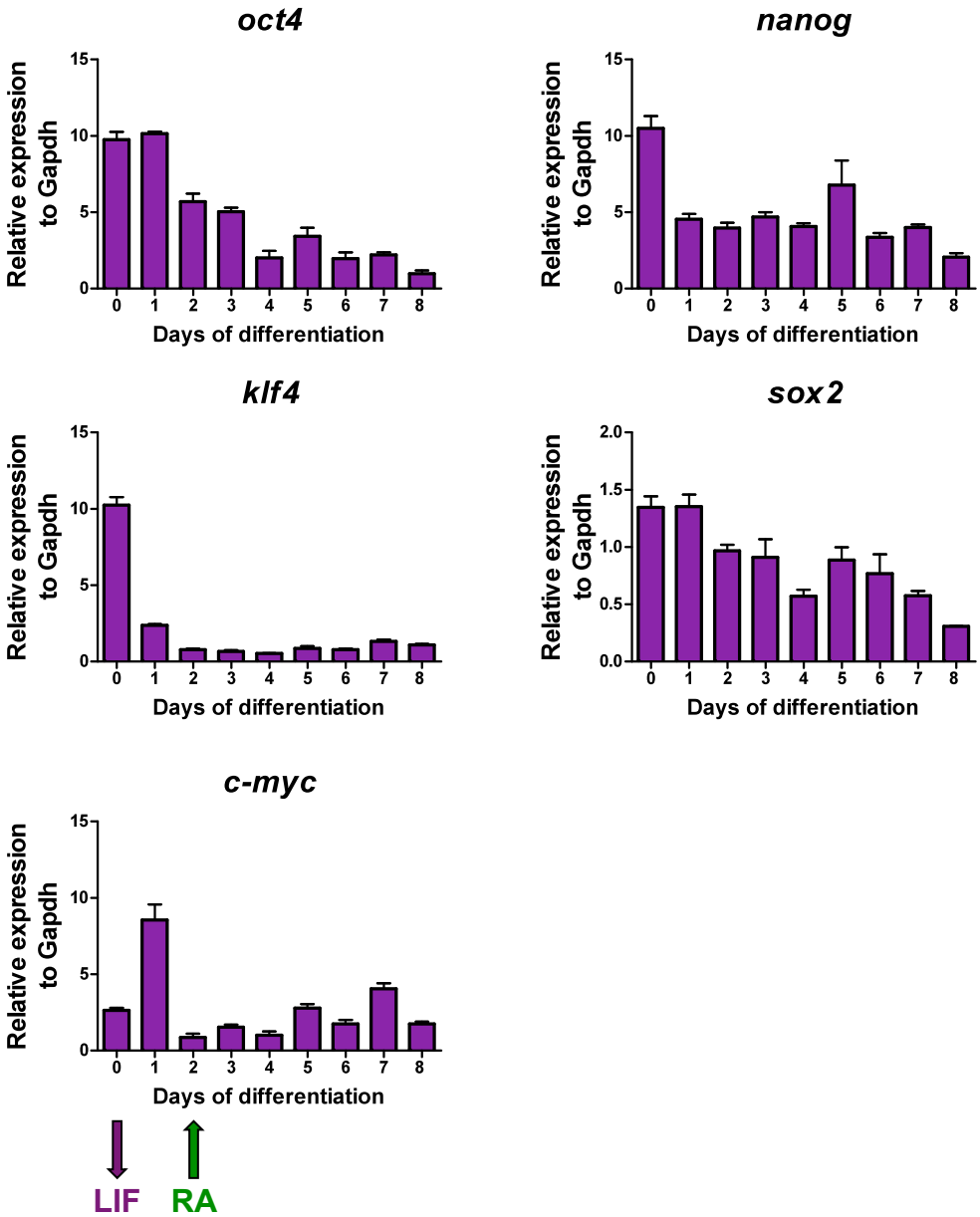
CCE/Rs were differentiated for 8 days using the standard protocol and observed through light microscopy. Images were taken at three different stages: undifferentiated cells, 24 hours and 7 days after LIF withdrawal.
Undifferentiated CCE/R were growing in colonies when one cell could not be distinguished from another. As expected, the differentiation induced the formation of floating embryonic bodies which were noticeably more compact at day 7.

3.2.2 Expression of pluripotency-associated genes and hoxb genes during differentiation of ES cells

RNA was extracted from differentiating CCE/R cultures every day from day 0 to day 8. The expression of five pluripotency-associated genes (*Oct4, Nanog, Klf4, Sox2* and *c-Myc*) was measured using qPCR. As expected, the expression of those genes dropped as cells were differentiating, with the exception of *c-myc*. One day after LIF withdrawal, the expression of *Klf4* was almost completely lost and remained absent until the end of the time course. In contrast, *Oct4* and *Sox2* didn't show this dramatic decrease, with the first sign of down-regulation visible only after 2 days of differentiation. Expression of the two genes then continued to drop progressively. Interestingly, half of the *Nanog* transcript was lost after 1 day of differentiation and stayed stable at this level until day 7 to finally drop again on day 8. Lastly, *c-myc* revealed a cycling profile, alternating between up- and down-regulation. Surprisingly, one day after LIF withdrawal, the gene was substantially up-regulated **(Figure 3.22)**.

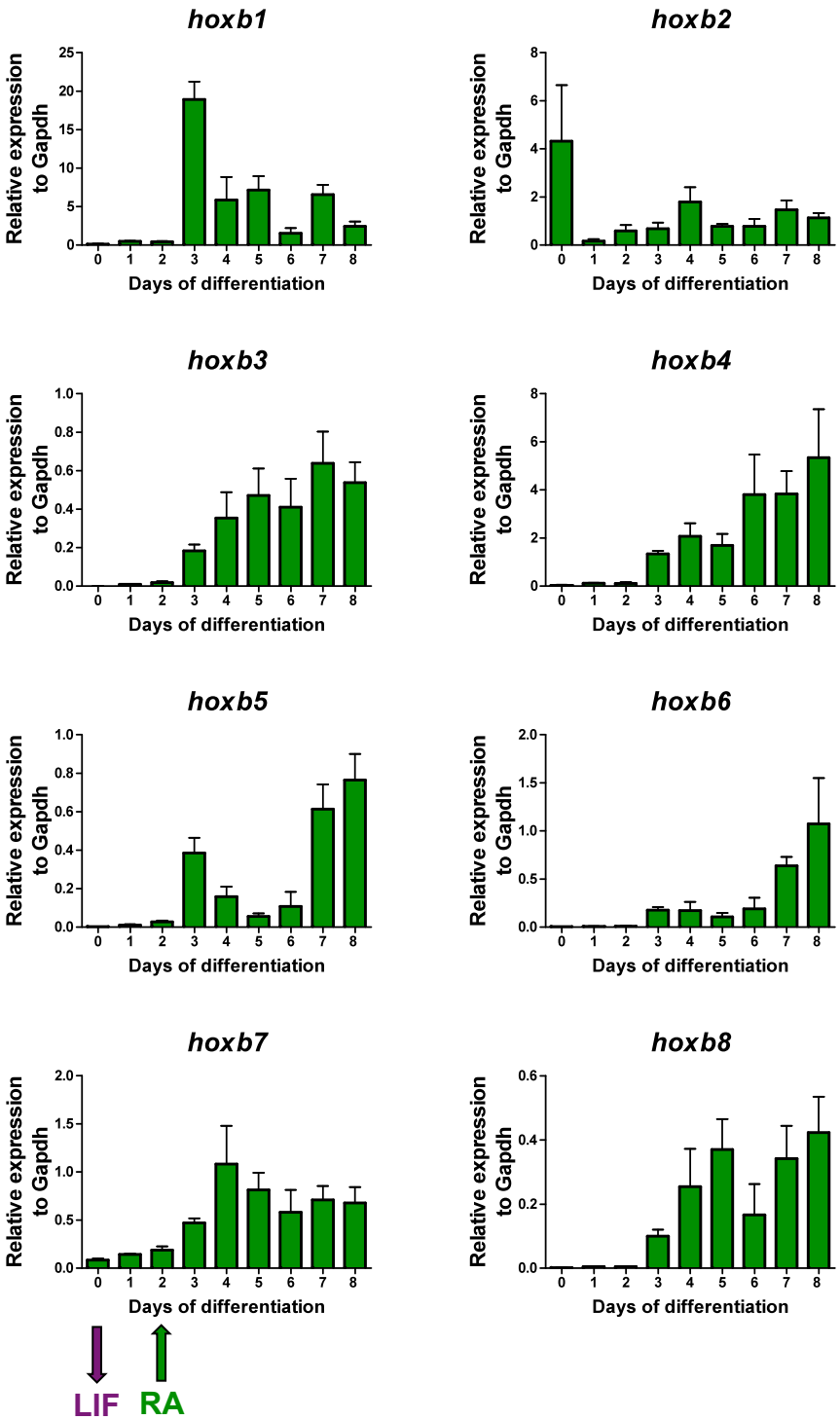
The expression of genes within the Hoxb cluster was then quantified, as RA treatment successfully induced the transcription of all *hoxb* genes we looked at. A typical profile of expression timing was observed, with for instance a peak of hoxb1 expression early during differentiation (day 3) and a peak of hoxb6 expression much later (day 8). Timing of induction generally correlated with the position of the gene in the cluster. However, *hoxb7* and *hoxb8* showed a particular feature, being activated earlier than predicted (day 4). In addition, while expression of *hoxb7* decreased between day 4 and 8, *hoxb8* was successively down- and up-regulated again **(Figure 3.23)**.

Figure 3.22 Pluripotency associated-gene expression during ES cell differentiation



CCE/R cells were induced to differentiate using a standard protocol over 8 days. RNA was extracted from differentiating cells every day and quantified by RT-qPCR using specific primers. Relative expression was calculated from standard curves. The expression of each gene was normalized against *Gapdh*. Error bars represent the standard error of the mean from three independent experiments.

Figure 3.23 *hoxb* gene expression during ES cell differentiation



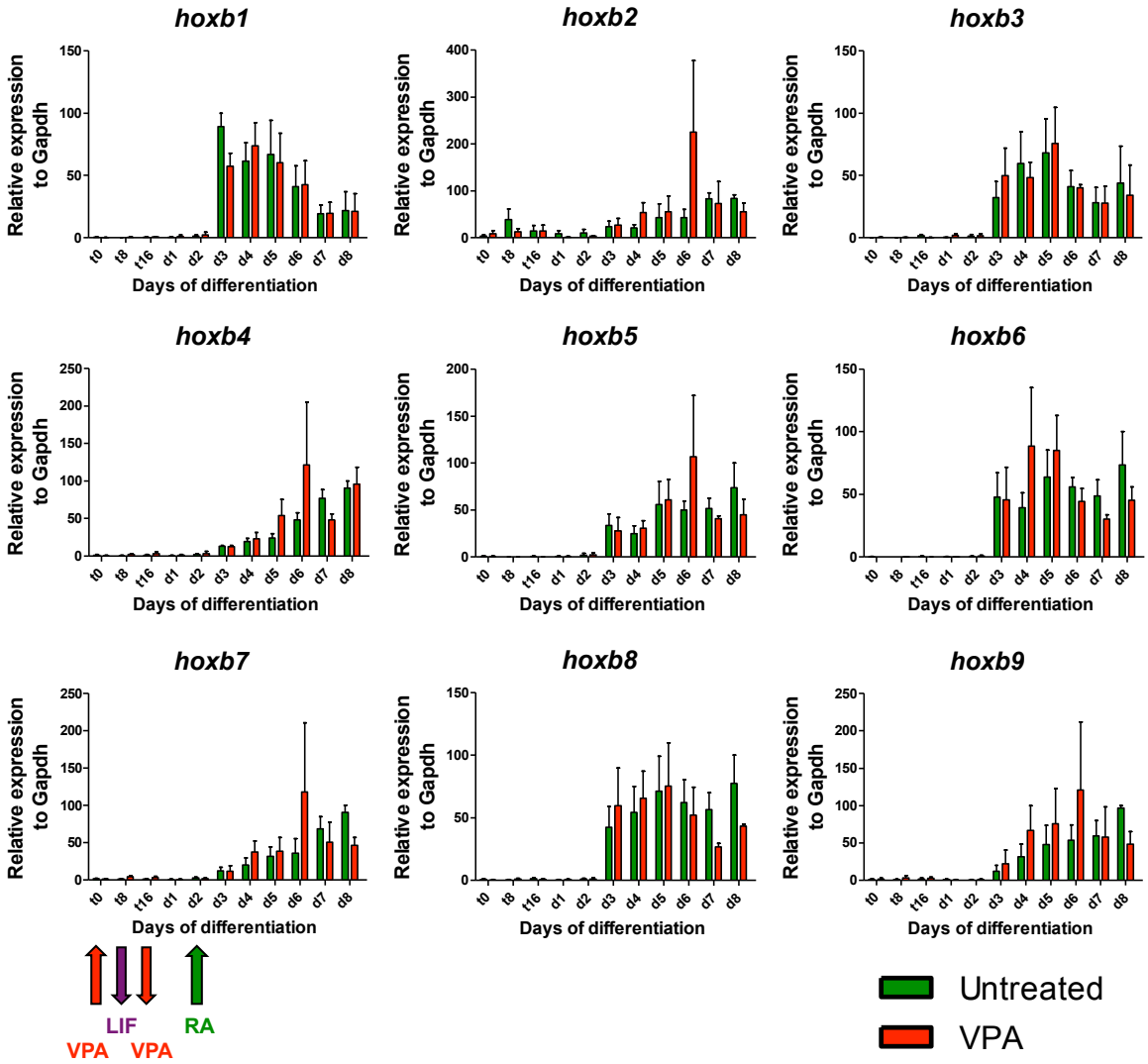
Cells were induced to differentiate using a standard protocol over 8 days. RNA was extracted from differentiating cells every days and quantified by RT-qPCR using specific primers. Relative expression was calculated from standard curves. The expression of each gene was normalized against *Gapdh*. Error bars represent the standard error of the mean from three independent experiments.

3.2.3 Valproic acid treatment during LIF withdrawal: Effect on Hoxb gene expression during differentiation

Using the same method, CCE/R cells were induced to differentiate. In addition, overlapping with the LIF withdrawal signal, histones were hyperacetylated by 1mM VPA treatment for 16 hours. RNA was extracted during the VPA treatment (T0hr, T8hrs and T16hrs) as well as every 24 hours from LIF withdrawal (T8hrs, d1-d8) **(Figure 3.20)**. Three independent experiments were carried out and results normalised using *Gapdh* and calculated as a percentage of the highest value in the untreated sample. As absolute values varied significantly between different experiments, this allowed the three sets of results to be merged and allowing visualisation of a general trend. The main observation was that the overall pattern of *hoxb* gene expression did not vary between untreated and treated samples. Sequential expression of genes in the cluster was observed throughout the differentiation time course, in the presence or the absence of VPA. Yet, a small effect was observed with over-expression of *hoxb* genes observed in the VPA-treated samples at certain time points. For instance in VPA-treated samples, the important peaks of *hoxb4* and *hoxb5* transcripts at day 6 or the more general over-expression of *hoxb9* throughout the time course **(Figure 3.24)**.

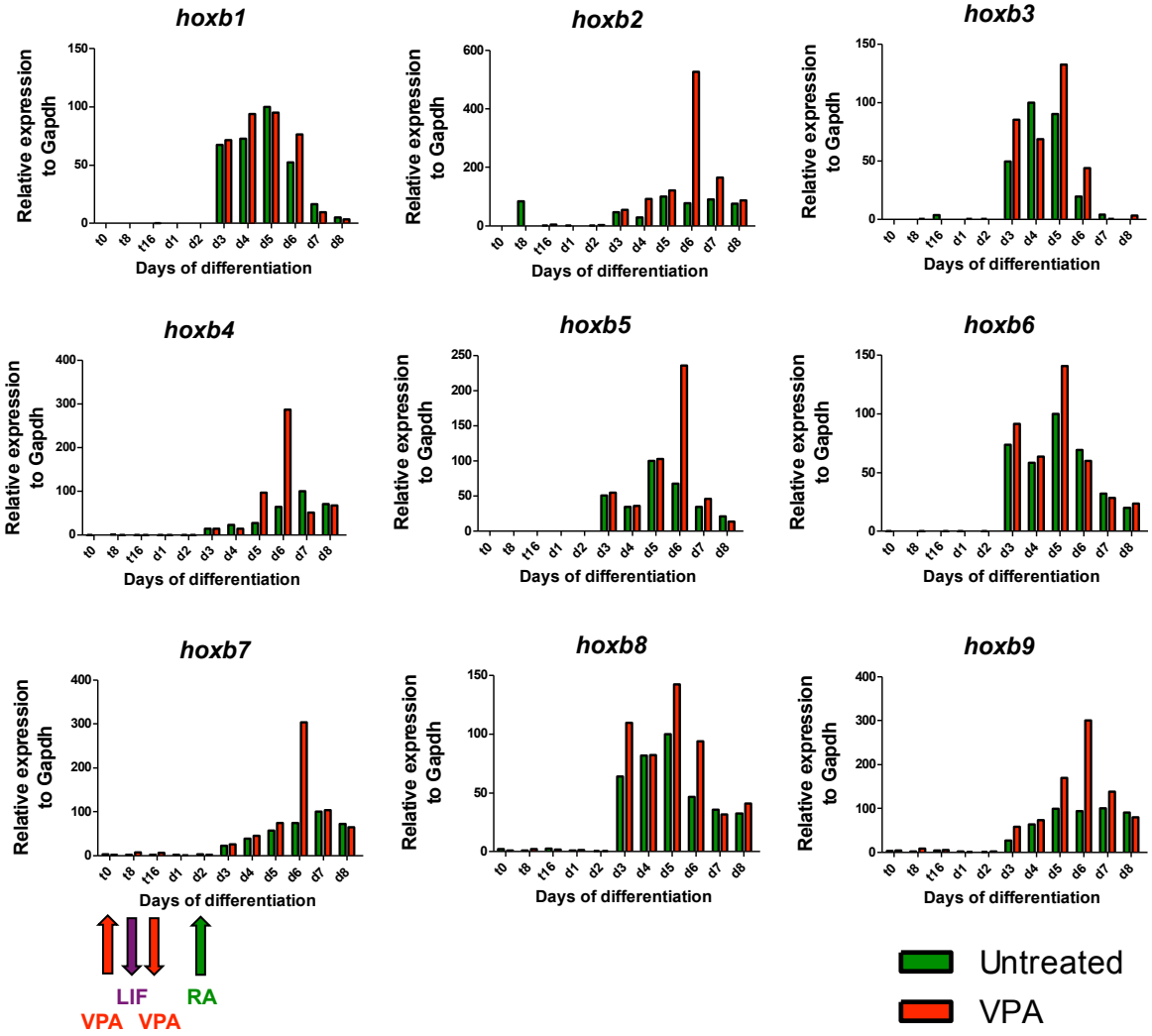
When experiments were analysed separately, the existence of hidden effects in the merged data was revealed. In fact, in two of the three independent experiments, peaks of expression of *hoxb* genes in VPA-treated samples were consistently larger and earlier than expression in control samples. Replicates 1 and 3 are presented as examples **(figure 3.25-26)**. Because those effects were observed on different days in the two experiments and because of the lack of effects in replicate 2, these effects

Figure 3.24 Effect of time-restricted VPA treatment on *hoxb* gene expression during differentiation (merged replicates)



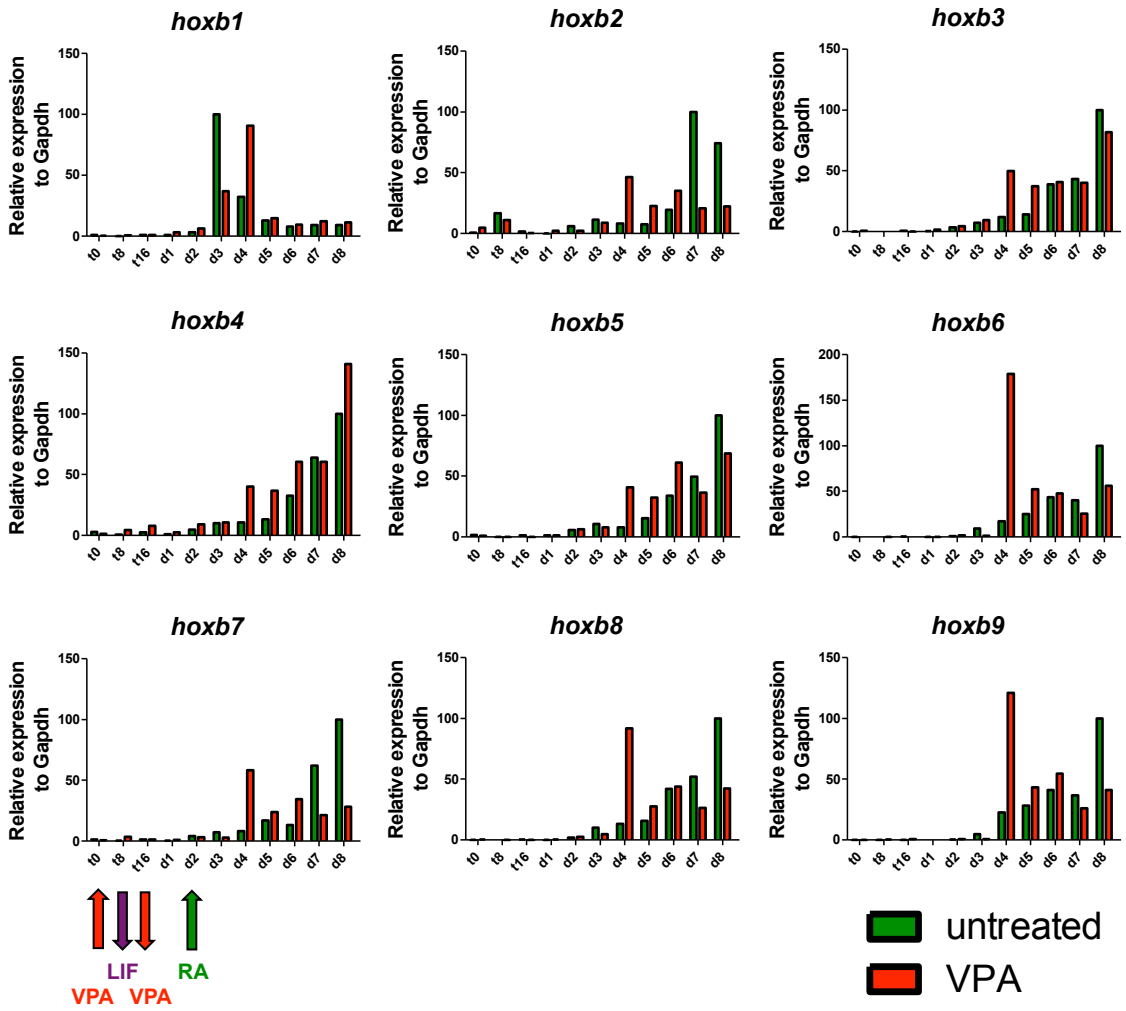
Control CCE/Rs were induced to differentiate using a standard protocol over 8 days. In parallel, CCE/Rs were treated with 1mM VPA for 16 hours while subjected to the same protocol. VPA was added 8 hours prior to LIF withdrawal and was washed out 16 hours later. RNA was extracted from control and treated differentiating cells every day and quantified by RT-qPCR using specific primers. Relative expression was calculated from standard curves. The expression of each gene was normalized against *Gapdh* and calculated as a percentage of the highest value in the untreated sample. Error bars represent the standard error of the mean from three independent experiments.

Figure 3.25 Effect of time-restricted VPA treatment on *hoxb* gene expression during differentiation (replicate 3)



Hoxb gene quantification in one of the three replicates from the experiment previously described in figure 3.25. Because merged data can sometimes hide the effects on individual experiments, we showed this replicate that could be representative of a real effect induced by VPA treatment.

Figure 3.26 Effect of time-restricted VPA treatment on *hoxb* gene expression during differentiation (replicate 1)



Hoxb gene quantification in one of the three replicates from the experiment previously described in figure 3.25. Because merged data can sometimes hide the effects on individual experiments, we showed this replicate that could be representative of a real effect induced by VPA treatment.

were not visible when data were merged. In order to validate one result or the other, it would be essential to do more experiments, possibly under more tightly controlled conditions. In fact, independent experiments were carried out at different times using different batches of unsynchronised cells.

So far, the inconsistency between replicates and the general trend observed when data were merged led us to conclude that the VPA treatment had no general effect on the timing of *hoxb* gene expression during differentiation.

3.2.4 summary

CCE/R were differentiated using cell detachment, LIF withdrawal and RA treatment. In order to investigate if histone modifications can predict gene transcription, we designed an experiment where the LIF withdrawal signal was supplemented by VPA treatment. In control conditions, it was possible to reproduce the specific pattern of *hoxb* gene expression when ES cells were differentiated with RA treatment. Addition of VPA to the culture medium at the start of the differentiation time course didn't globally affect *hoxb* gene expression timing. However, analysis of individual experiments suggested the hyperacetylation of histones at the beginning of the time course could result in the premature expression of *hoxb* genes during differentiation.

3.3 Role of polycomb silencing mechanism in the response to VPA treatment: the hoxb gene cluster case

3.3.1 Introduction

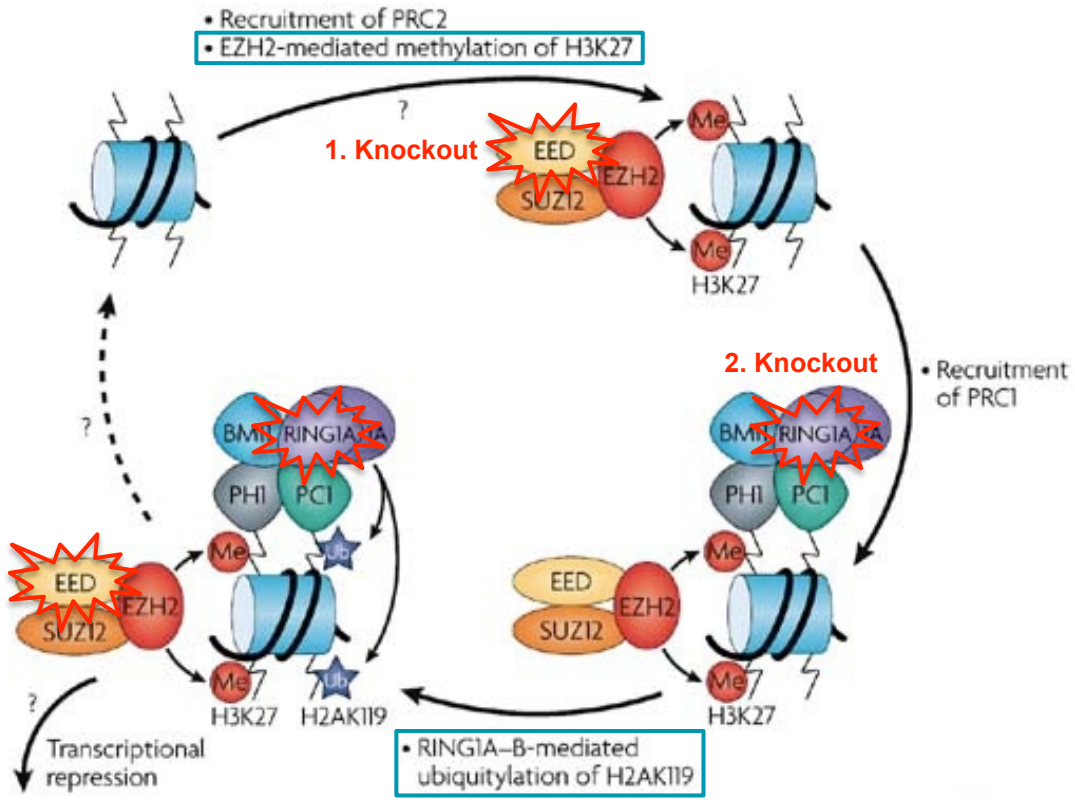
When ES cells were treated with VPA, we observed that the effect on global gene expression was not as extensive as expected. In fact, despite the massive global hyperacetylation of histones, only 10% of genes examined using microarrays showed a change in expression. In addition, although acetylation of histone has been associated with active transcription, almost equal proportions of genes were up- and down-regulated by VPA treatment. This small effect of VPA on transcription raised the possibility that some mechanisms protect genes against hyperacetylation. HDACi are naturally present in the environment, as products of the gut microflora, or as dietary components, and this could justify the need for processes that maintain normal gene expression levels within a cell in the presence of HDACi.

One could imagine that a modification on its own does not drive transcription and hyperacetylation could be buffered by “cross-talk” with other epigenetic events. Here we decided to assess whether there is a role for polycomb-deposited marks in the response to hyperacetylation. The polycomb proteins assemble in two complexes (PRC1 and PRC2) that put in place two histone modifications, tri-methylation at H3K27 (H3K27me3) and ubiquitination at H2AK119 (H2AK119ub), associated with repressed transcription. Polycomb complexes regulate an important number of developmental genes, playing a crucial role in the maintenance of pluripotency and cell fate decision in ES cells (Boyer et al., 2006). The homeobox (hox) genes are known targets of polycomb group proteins. Hox genes are not expressed in pluripotent embryonic stem cells and are activated in a very specific manner during

differentiation in order to specify the antero-posterior axis of the embryo. Polycomb and Trithorax proteins, that respectively repress or activate gene expression by the deposition of specific marks, play a key role in the regulation of hox genes in ES cells and in differentiated cells (Bernstein et al., 2006).

We decided to quantify *hoxb* gene expression under VPA treatment in an ES cell line deficient for both the PRC1 component *Ring1b* and the PRC2 component *Eed* (Leeb et al., 2010). If *hoxb* genes were up-regulated in Eed/Ring1b double knock-out (dko) cells but not in the corresponding wild type (wt) cells, the role of polycomb in the maintenance of gene silencing despite the hyperacetylation of histones would be highlighted. After we characterised the dko cell line grown under our conditions, wt and dko cells were treated with VPA to assess the impact on global histone modifications and *hoxb* gene activity **(Figure 3.27)**.

Figure 3.27 Model system #3:
Role of polycomb group proteins and their associated histone marks in the response to the HDACi VPA



3. Effect of the knockout
4. Effect of the knockout combined with VPA treatment.

Previous studies have shown that polycomb proteins maintain the silencing of *hoxb* genes in ES cells. Expression of those genes was measure in Wild type (wt) ES cells and in *Eed-Ring1b* double knockout (dko) ES cells. *Eed* and *Ring1b* code for two major proteins, members of Polycomb complexes PCR2 and PCR1 that respectively put in place the silencing marks H3K27me3 and H2AK119ub. Wt and dko cells were treated with VPA and *hoxb* gene expression was compared in untreated and VPA-treated cells.
Figure adapted from (Spivakov and Fisher, 2007).

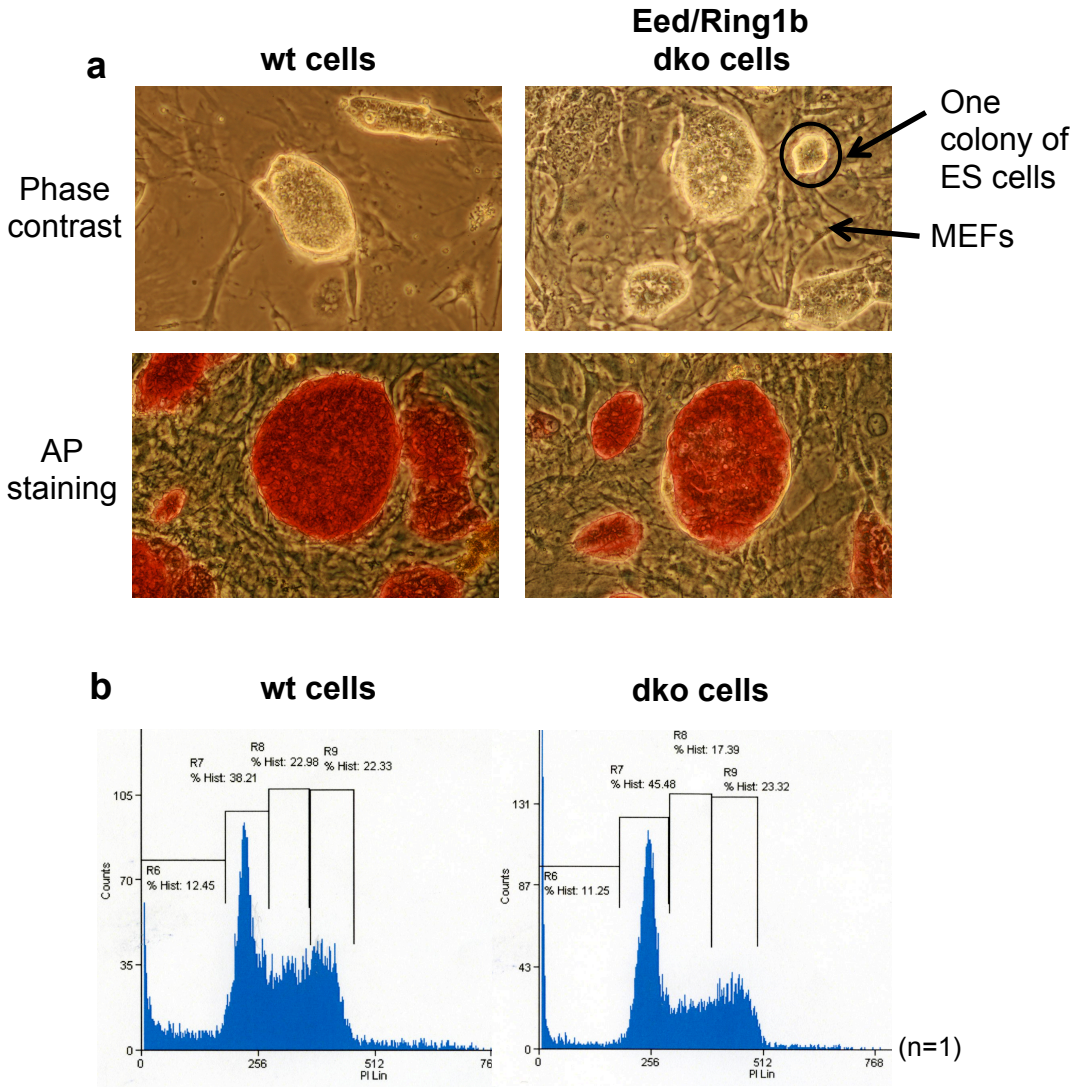
3.3.2 Characterisation of the double knockout cell line

3.3.2.1 Cell morphology and cell cycle

Cell morphology is an important indicator of ES cell pluripotency. Using light microscopy, we looked at wt and dko cells grown on a layer of MEF feeder cells. Both cell lines showed typical growth morphology with cells forming colonies. In addition, both cell lines strongly stained for alkaline phosphatase activity, a marker of pluripotent (undifferentiated) cells. As expected, ES cells appeared red while MEFs remained unstained, which validated the use of the alkaline phosphatase marker for ES cells **(Figure 3.28.a)**.

We then assessed cell cycle profiles of wt and dko cells using flow cytometric analysis of cells stained with propidium iodide (PI). 50000 cells on average were analysed and the intensity of the PI signal for each cell was measured. In a wt cell culture, 38% of cells were in the G1-phase and 23% were in the S-phase. In the dko cell culture those percentages changed to 45% and 17% respectively. This could be the result of technical variation, as demarcation of each population on the histogram is inaccurate for FACS analysis. But this result was consistent with observations that dko cells grew slower than wt cells in culture. These variations possibly highlighted the fragile nature of dko cells and their tendency to easily differentiate. However, those changes were considered very small and the dko cell cycle profile was still viewed as a typical ES cell cycle profile **(Figure 3.28.b)**.

Figure 3.28 Effect of a Eed/Ring1b double knockout on ES cell morphology, alkaline phosphatase activity and cell cycle



a, Wt and Eed/Ring1b dko cell morphologies were observed using light microscopy. Both cell lines were grown on a layer of mouse embryonic fibroblasts (MEF) feeder cells. Dko cells required double the amount of MEFs in order to maintain pluripotency. Cells were stained for alkaline phosphatase activity (AP), a marker of pluripotency. Pluripotent, AP-positive colonies appeared red.

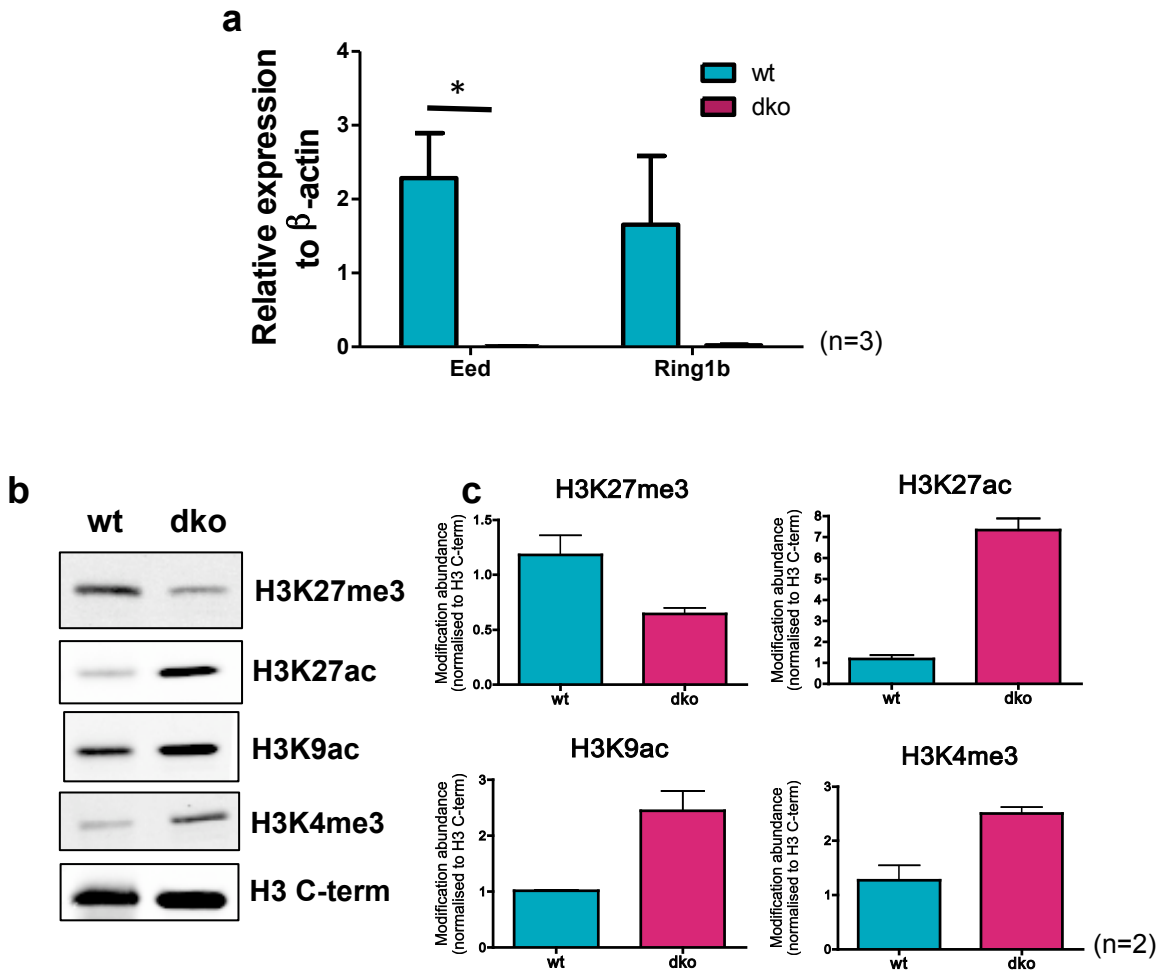
b, Wt and dko cells were stained with propidium iodide (PI), a DNA intercalating dye. The intensity of the signal from each cell was then analysed using flow cytometry in order to obtain the percentage of cells in each phase of the cell cycle.

3.3.2.2 Global histone modifications

RNA was extracted from wt and dko cells and the expression of *Eed* and *Ring1b* was quantified by RT-qPCR. As expected, the expression of those genes was lost in dko cells **(Figure 3.29.a)**. Following the verification of gene knockouts, we investigated the effect on global histone modification levels. Histones were extracted from wt and dko cells and levels of four histone marks were quantified using western blotting **(Figure 3.29.b)**. The secondary antibody was conjugated with a fluorescent dye and the signal was quantified with the LICOR Odyssey scanner. Each value was normalized using an antibody raised against H3 C-terminal domain **(Figure 3.29.c)**. Despite the loss of *Eed* and *Ring1b* transcription, the H3K27me3 mark was not fully abolished in dko cells. The H3K27me3 enrichment was undoubtedly reduced in dko cells (≈ 2 fold) but still present. This intriguing observation could simply reflect a persisting proportion of MEFs after ES cell purification, as a small quantity of MEFs could generate the H3K27me3 signal. On the other hand, the ES cell culture we examined could be heterogeneous with wt and dko cells growing together, though this seems unlikely given the absence of *Eed* and *Ring1b* transcripts in dko cells. Finally, it is possible that another enzyme compensated for the loss of *Eed* and deposited the H3K27me3 mark.

We decided to measure H3K27ac as well because this trithorax-dependant mark is understood to antagonize the polycomb-mediated silencing (Tie et al., 2009). Indeed, the decrease of H3K27me3 in dko cells was concomitant with a major increase of H3K27ac (≈ 7 fold) when compared with wt cells. The trithorax-dependant mark H3K4me3 was also increased in dko cells (≈ 2 fold), as well as H3K9ac (≈ 2 fold).

Figure 3.29 Measure of Eed/ring1b expression and global histone modification in wt and dko cells

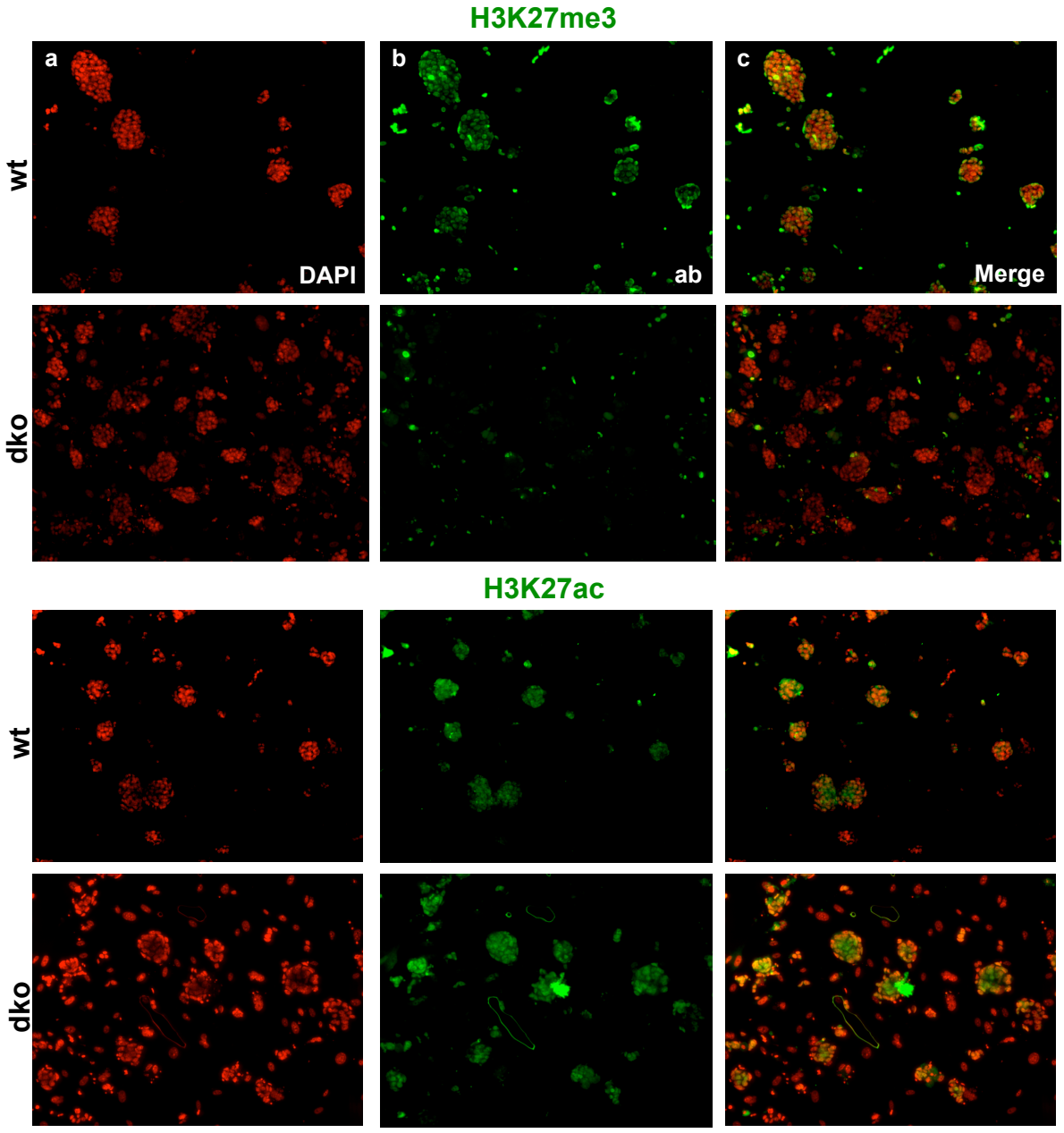


a, RNA was extracted from wt and Eed/Ring1b dko cells and quantified by RT-qPCR using specific primers. The relative expression of *Eed* and *Ring1b* genes was calculated from standard curve and normalized against β-actin. Parametric two-tailed unpaired t-test were processed with the GraphPad software. Error bars represent the standard error of the mean from three independent experiments.
b, Histones were extracted from wt and Eed/Ring1b dko cells. Changes in global levels of H3K27me3, H3K27ac, H3K9ac and H3K4me3 were assessed by western blotting. H3 C-term staining was used as a loading control.
c, Quantitative presentation of modification abundance. H3 C-terminal staining was used to normalize values. Error bars represent the standard error of the mean from two independent experiments.

Once again, the analysis of global histone modification levels suggested cross-talk and the interdependence of different marks.

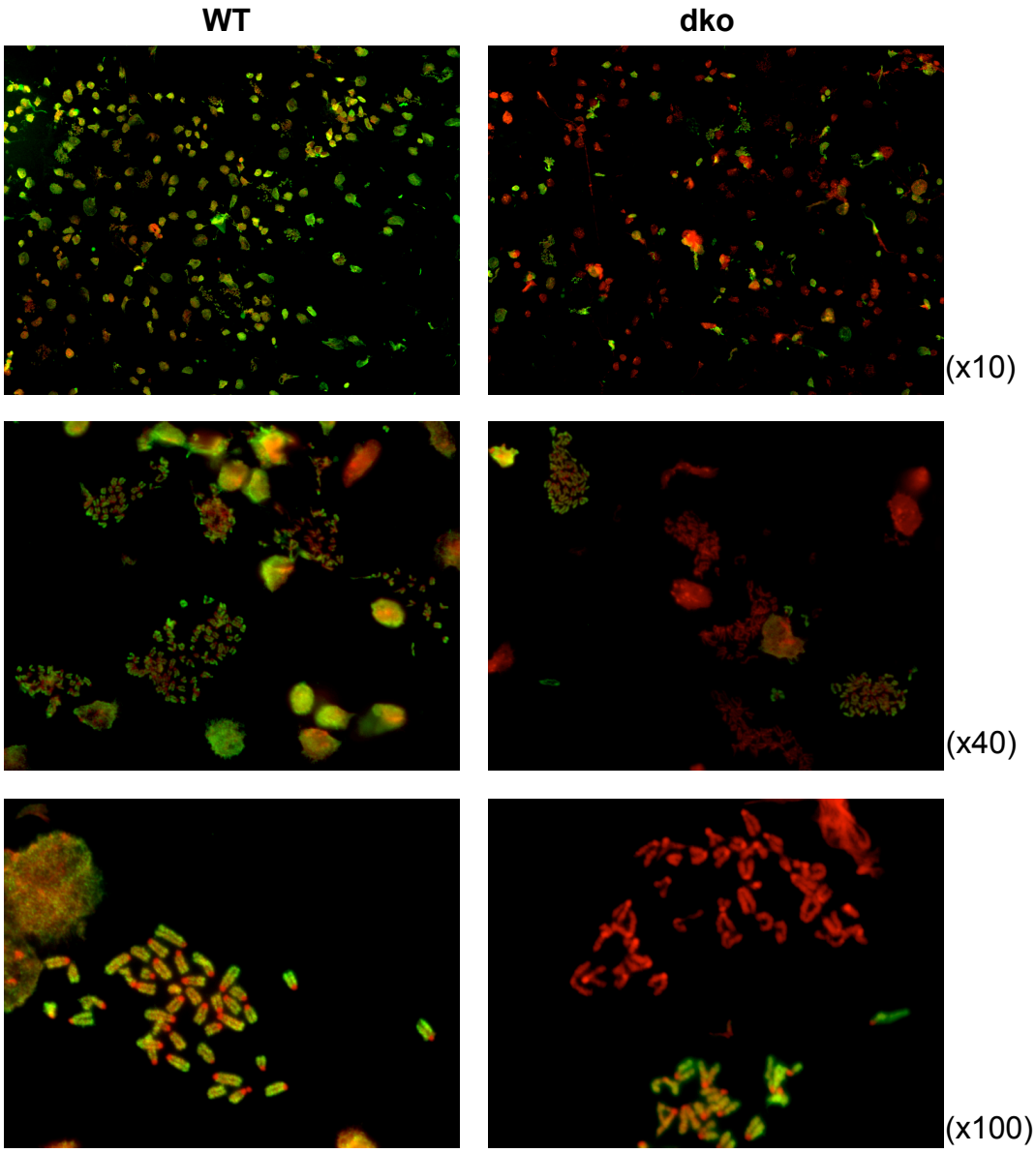
In order to investigate more the residual H3K27me3 modification in dko cells, immuno-fluorescent staining was used to examine the location of this histone mark. Firstly, colonies of wt and dko cells grown on coverslips were directly fixed and stained with antibodies raised against H3K27me3 and H3K27ac. In addition, DAPI staining allowed the visualisation of nuclei. As expected, H3K27ac was present in wt and dko cells. In contrast, while wt cells were enriched in the H3K27me3 mark, dko cells appeared depleted **(Figure 3.30)**. However, the resolution of the technique was not very high and quantification was not possible. We therefore decided to extend this study by staining metaphase chromosomes. Wt and dko cells were treated with colcemid, which induces an arrest in metaphase, and trypsinised to obtain a single cell suspension. Cells were then fixed to slides, and stained for H3K27me3. In wt cells, all cells and each of the chromosome spreads showed strong green staining meaning that the histone mark H3K27me3 was present. Surprisingly in dko cells the staining was heterogeneous, with 68.4% of cells appearing completely red (DAPI staining only, average count from 6 images, n=2) while 31.6% of cells stained for H3K27me3. Chromosome spreads showed the same pattern with some spreads entirely red whereas others were identical to wt **(Figure 3.31)**. This observation almost eliminated the possibility of a redundant mechanism for H3K27me3 deposition, as in this case, staining would be weaker but homogeneous. Positive cells for H3K27me3 could therefore be MEFs or wt cells. Prior to incubation with ES cells, MEFs were irradiated to stop their division. This leads to DNA damage that was not visible on the dko stained spreads. In addition, the proportion of cells

Figure 3.30 Immunofluorencent labelling of wt and Eed/Ring1b dko ES cell colonies for H3K27ac and H3K27me3



Wt and dko cells were grown on coverslips in a 24-well plate. Cells were then directly fixed and immunostained with antibodies raised against the two histone modifications H3K27me3 and H3K27ac. Images were taken using a fluorescent microscope. The red staining corresponds to the DAPI signal, a molecule that strongly fluoresces when bound to DNA (**a**). The green staining (FITC) reveals the presence of the marks (**b**). Finally, both signals were merged in order to detect the enriched nuclei (**c**).

Figure 3.31 Immunofluorencent labelling of wt and dko mataphase chromosomes for the H3K27me3 mark



In order to block wt and dko cells in metaphase, Colcemid was added into the culture medium for 3 hours prior harvesting. Cells were then fixed and immunostained with an antibody raised against the histone modification H3K27me3. The red staining corresponds to the DAPI signal, a molecule that strongly fluoresces when bound to DNA. The green staining corresponds to the secondary antibody that reveals the presence of the mark. Both signals where merged and images were taken using a fluorescent microscope at different magnifications.

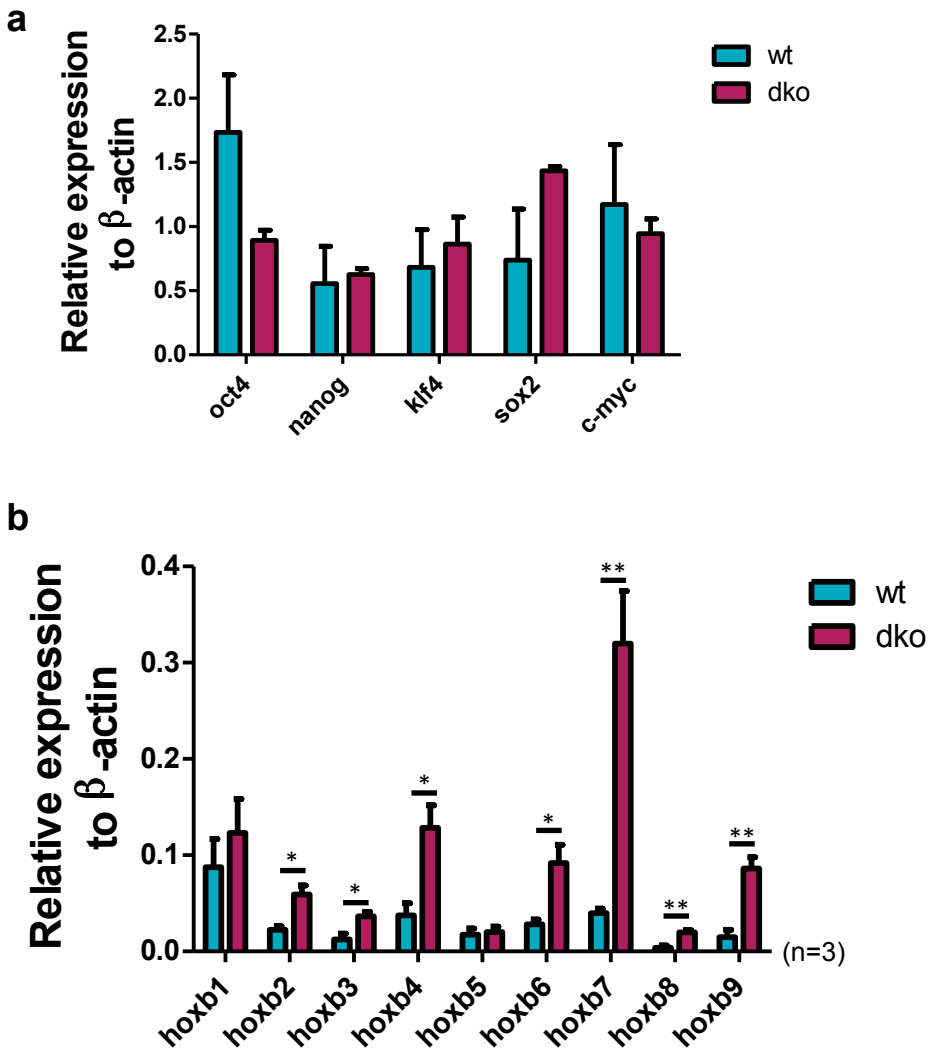
and spreads positive for H3K27me3 appeared too large to reflect MEF contamination. In fact, metaphase chromosomes come from dividing cells and yet MEFs were not cycling due to the irradiation treatment. For these reasons, even if the other possibilities were not fully excluded, we concluded that our dko cell population wasn't pure and contained a small proportion of wt cells.

3.3.2.3 Gene expression

Finally, we decided to quantify pluripotency-associated genes and *hoxb* genes in wt and dko cells. RNA was extracted from wt and dko cells and the expression of specific genes was quantified using RT-qPCR. *Nanog*, *Klf4* and *c-Myc* were expressed at the same level in dko and wt cells. On the other hand, *Oct4* was down regulated in the dko cells while *Sox2* was slightly up-regulated. However, those differences were not statistically significant **(Figure 3.32.a)**.

The expression of *hoxb* genes in wt and dko cells was then quantified. With the loss of polycomb silencing marks, most of the genes in the cluster were up-regulated in dko cells. However, the fold changes were not uniform across the cluster. *Hoxb7* showed the most important increase (≈ 6 fold) followed by *hoxb4*, *hoxb6* and *hoxb9* (≈ 3.5 to 4.5 fold). The effect of the knockout appeared less strong on *hoxb2*, *hoxb3* and *hoxb8*, whereas *hoxb1* and *hoxb5* were not affected by the knockout **(Figure 3.32.b)**.

Figure 3.32 Effect of the double knockout on the expression of pluripotency associated genes and *hoxb* genes



RNA was extracted from wt and dko cells and retrotranscribed into cDNA before quantification. Expression of specific genes was measured using qPCR and normalized to β-actin. Parametric two-tailed unpaired t-tests were processed with the GraphPad software. Error bars represent the standard error of the mean from three independent experiments.

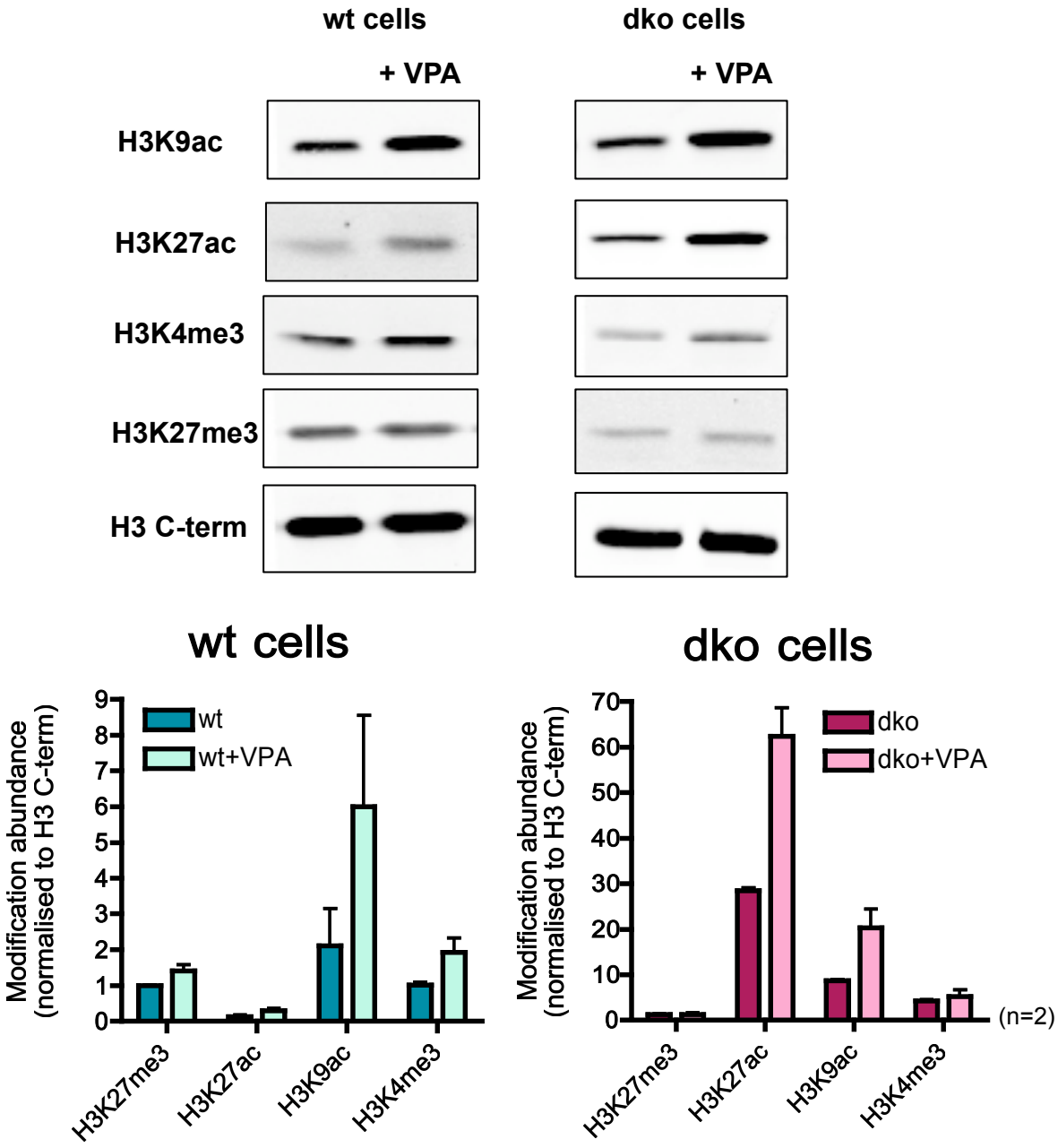
3.3.3 Effect of valproic acid treatment on global histone modification levels in wt and dko cells

Histones were extracted from wt and dko cells before and after an 8 hours exposure to 1mM VPA. Global levels of H3K9ac, H3K27ac, H3K4me3 and H3K27me3 were obtained by western blot analyses and normalised to the H3 C-term signal. As expected the enrichment of H3K9ac and H3K27ac increased following VPA treatment, in both wt and dko cells. For those modifications, fold changes were similar in the two cell lines. As already seen with the CCE/R cell line, VPA also induced a small increased in H3K4me3, but this was a larger effect in wt cells than in the double knockout. Finally, H3K27me3 abundance remained unaffected by VPA treatment **(Figure 3.33)**.

3.3.4 Effect of valproic acid treatment on gene expression in wt and dko cells

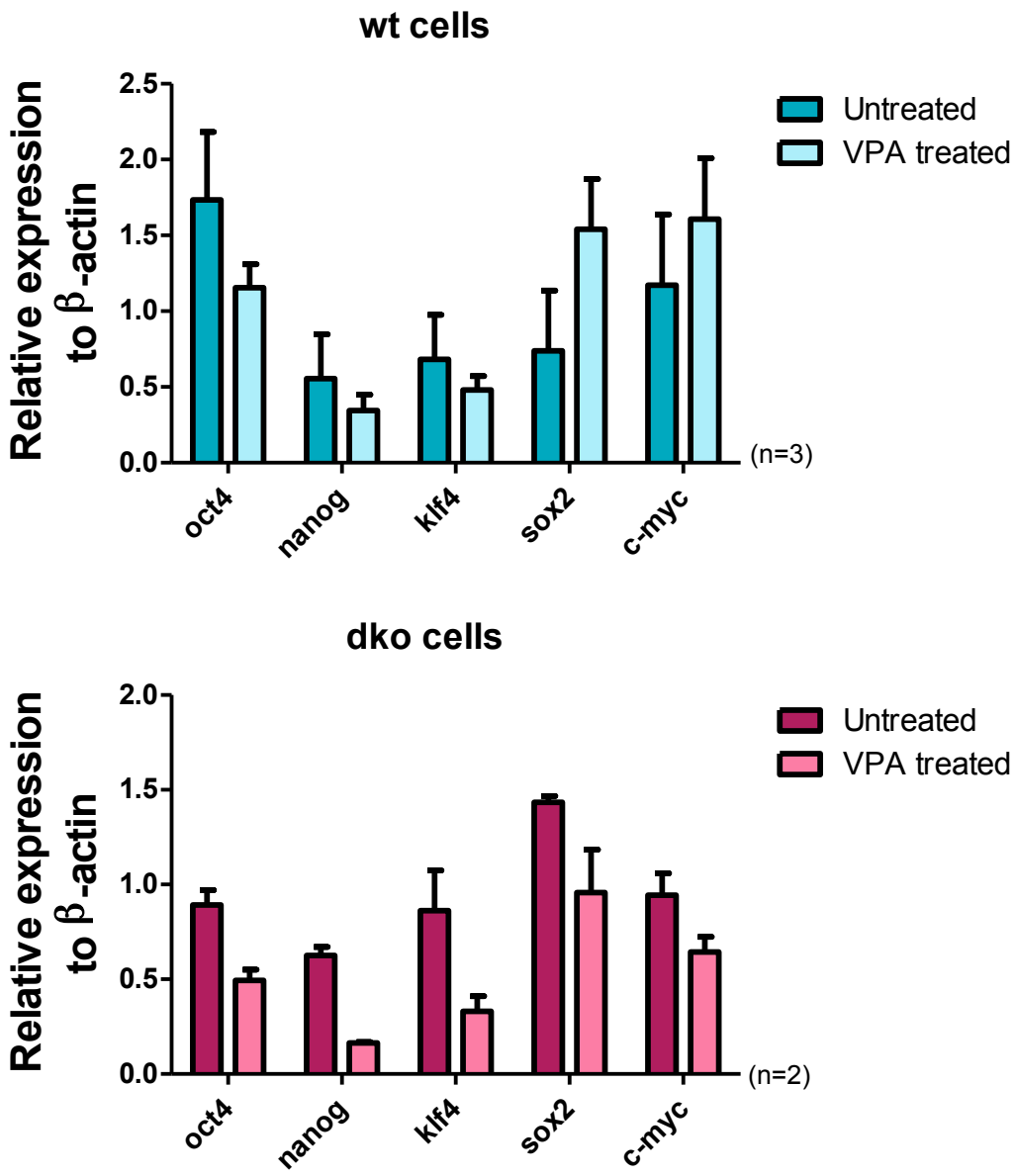
RNA was extracted from wt and dko cells that had been treated with 1mM VPA for 8 hours. VPA is known to affect pluripotency in ES cells. We showed with the CCE/R cell line that a 1mM VPA treatment for 8 hours induced a reduction of *Nanog*, *Klf4* and in some cases *Oct4* expression. *c-Myc* was not affected by the treatment and *Sox2* was up-regulated. Although the results were not statistically significant, the wt cell line was similarly affected by VPA - transcription of *Oct4*, *Nanog* and *Klf4* was reduced after treatment and up-regulation of *Sox2* was observed. In the dko cell line, all pluripotency-associated genes were down-regulated, especially *Nanog* and *klf4*. This result was consistent with the fragile nature of the dko cells and their tendency to easily differentiate **(Figure 3.34)**.

Figure 3.33 Effect of VPA treatment on global histone modifications in wt and Eed/Ring1b dko ES cells



Histones were extracted from wt and Eed/Ring1b dko cells before and after VPA treatment for 8 hours. The effect on levels of global histone modifications was assessed using western blotting and normalised using H3 C-terminal staining. Error bars represent the standard error of the mean from two independent experiments.

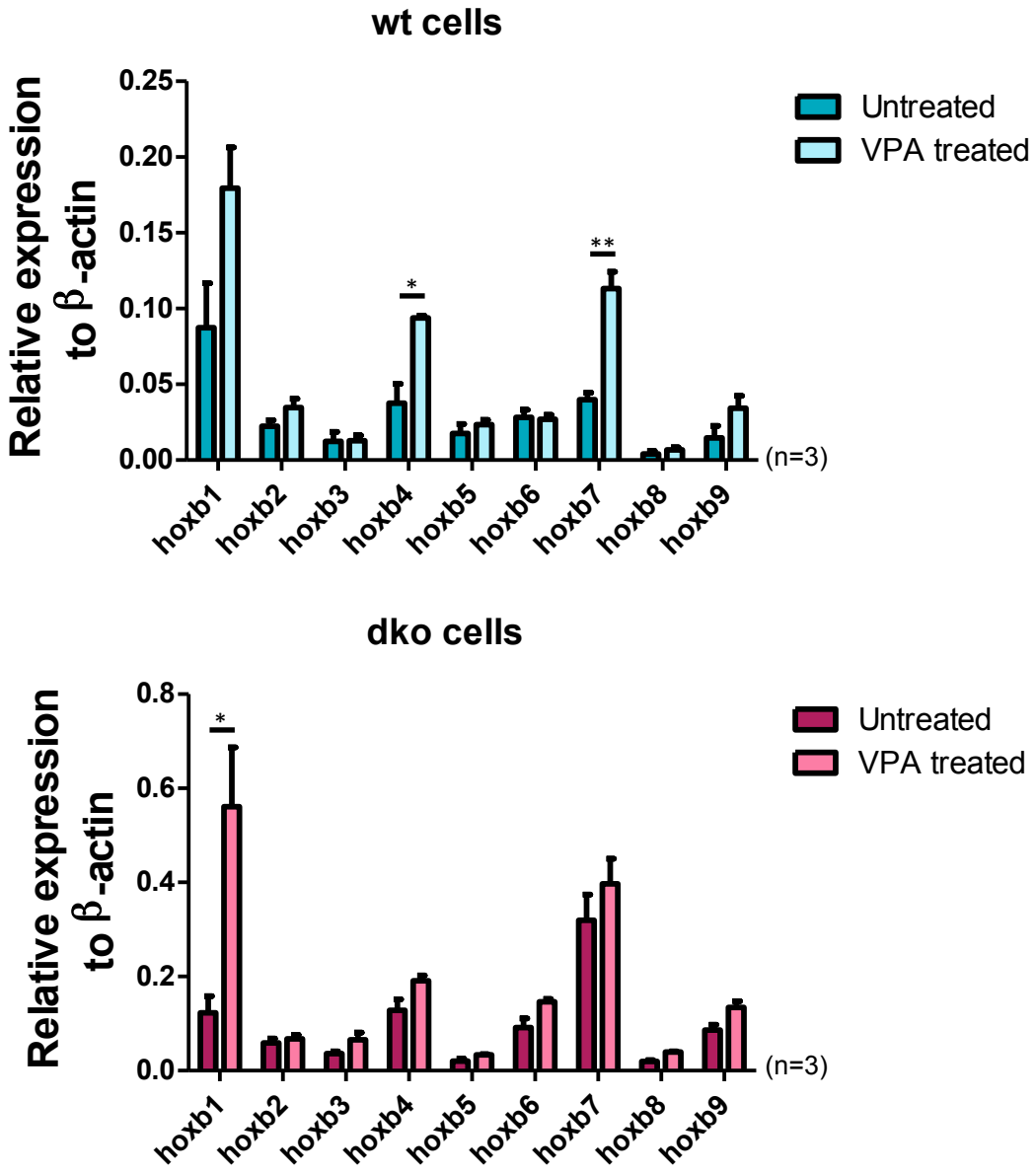
Figure 3.34 Effect of VPA treatment on the expression of pluripotent-associated genes in wt and Eed/Ring1b dko ES cells



Wt and dko cells were treated with 1mM VPA for 8 hours. RNA was extracted and retrotranscribed into cDNA before quantification by qPCR. The relative expression of pluripotency-associated genes was measured using a standard curve and normalized to β-actin. Error bars represent the standard error of the mean.

The expression of genes within the *hoxb* cluster was quantified in wt and dko cells before and after VPA treatment. Very interestingly, not all the genes responded to the treatment in a similar manner. Futhermore, the changes in pattern of gene expression induced by the treatment were different in wt cells compared to dko cells. In wt cells, transcription of *hoxb1*, *hoxb4*, *hoxb7* and *hoxb9* was increased by more than 2-fold. On the other hand, the expression of *hoxb2*, *hoxb3*, *hoxb5*, *hoxb6* and *hoxb8* was unaffected or showed a very small increase after VPA exposure. In dko cells, the only gene significantly up-regulated by VPA treatment was *hoxb1*. Very small increases were visible on the other genes in the cluster but all were non significant. We therefore conclude that, with the exception of *hoxb1*, VPA had no effect on the transcription of *hoxb* genes in dko cells **(Figure 3.35)**.

Figure 3.35 Effect of VPA treatment on the expression of *hoxb* genes in wt and Eed/Ring1b dko ES cells



WT and dko cells were treated with 1mM VPA for 8 hours. RNA was extracted and retrotranscribed into cDNA before quantification by qPCR. The relative expression of *hoxb* genes was measured using a standard curve and normalized to β-actin. Parametric two-tailed unpaired t-tests were processed with the GraphPad software. Error bars represent the standard error of the mean of three independent experiments.

3.3.5 Effect of valproic acid treatment on *hoxb* gene expression in CCE/R cells

In order to confirm these observations made with wt cells, we decided to quantify the expression of *hoxb* genes in the CCE/R ES cell line before and after an identical VPA treatment. Apart from *hoxb2,* that showed high expression compared to the other genes, these results were similar to those obtained with the wt cell line, with VPA inducing increased hoxb1, hoxb4 and hoxb7 transcripts. Meanwhile, the expression of other *hoxb* genes in the cluster remained stable **(Figure 3.36)**.

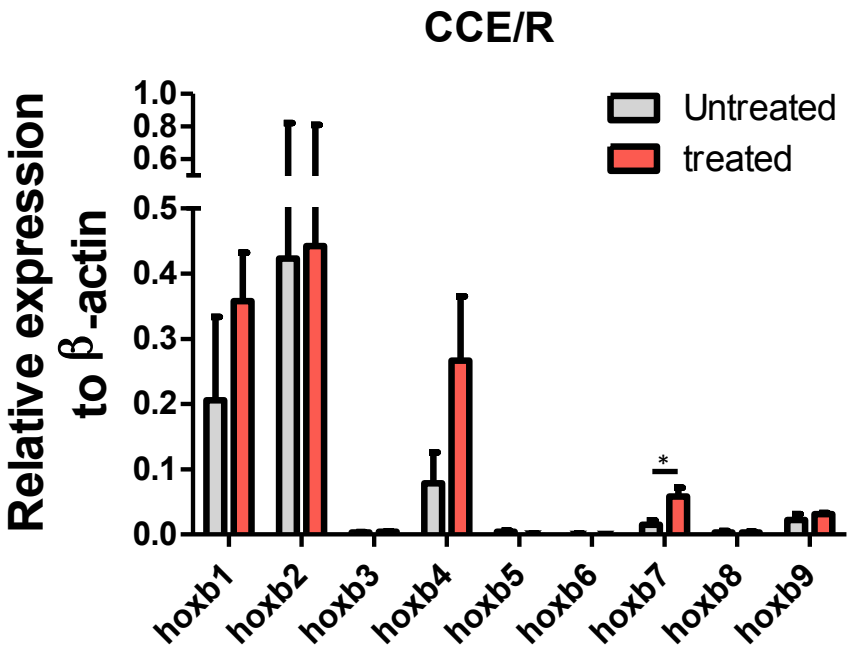
3.3.6 Effect of valproic acid treatment on histone modifications at hoxb genes in CCE/R cells

Hoxb1, *4* and *7* were consistently up-regulated after VPA treatment in two different wild-type ES cell lines. This result suggested that regulation by histone modifications was not uniform across the hoxb cluster. For this reason, the abundance of histone modifications was assessed at the three affected genes (*hoxb1*, *4* and *7*), and at two unaffected genes (*hoxb2* and *5*). The enrichment of marks was measured before and after eight hours VPA treatment at the promoter, the TSS, and in the first exon of each gene **(Figure 3.37)**. Levels of three activating marks (H3K9ac, H4K8ac and H3K4me3) and one silencing mark (H3K27me3) were examined.

As expected in untreated cells, levels of H3K27me3 were very high at all hoxb genes. Although down-regulation was observed at certain sites after VPA treatment, these effects were very small and it was concluded that H3K27me3 was mainly unaffected by the drug. Surprisingly however, and with the exception of *hoxb4* and *7*, hoxb genes appeared depleted in the second bivalent mark H3K4me3, as well as in acetyl marks in untreated cells. H3K4me3 and H4K8ac remained unaffected by VPA

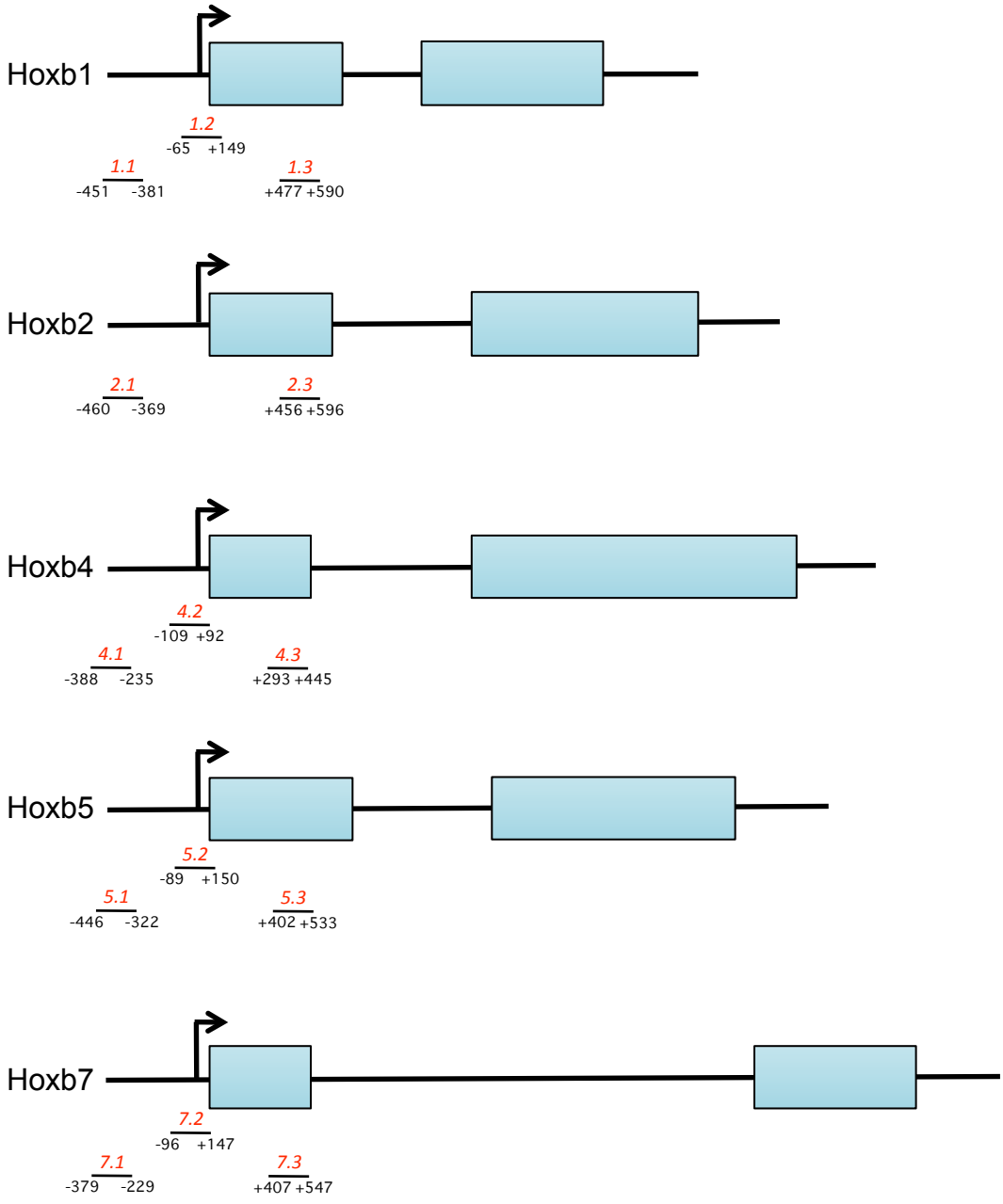
treatment, whereas levels of H3K9ac increased at all hoxb genes, regardless of transcriptional activities **(Figure 3.38)**.

Figure 3.36 Effect of VPA treatment on the expression of *hoxb* genes in CCE/R ES cells



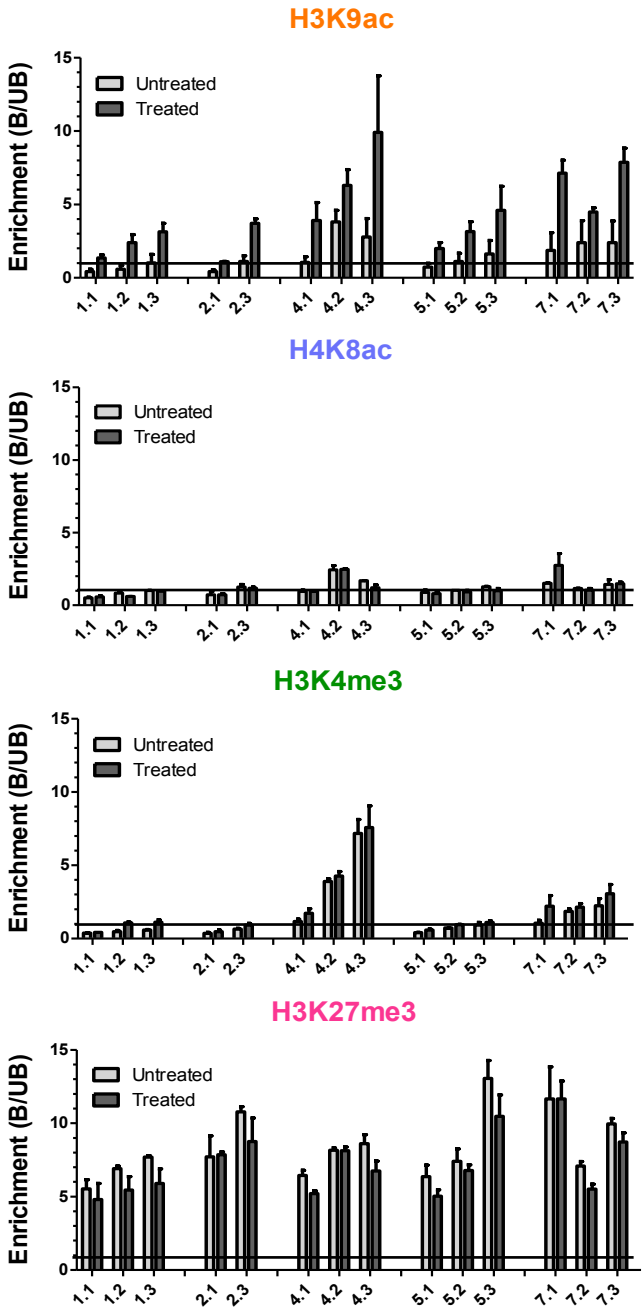
CCE/R cells were treated with 1mM VPA for 8 hours. RNA was extracted and retrotranscribed into cDNA before quantification by qPCR. The relative expression of *hoxb* genes was measured using standard curve and normalized to the β-actin. Parametric two-tailed unpaired t-tests were processed with the GraphPad software. Error bars represent the standard error of the mean of three independent experiments.

Figure 3.37 Distribution of primer sets used for qPCR analysis over *hoxb* genes



For each *hoxb* gene, the arrow corresponds to the transcription start site and boxes indicate exons. Numbers below each primer set indicate the base pairs surrounding the amplified region.

Figure 3.38 Distribution of histone modifications on hoxb genes before and after VPA treatment



CCE/Rs were treated with 1mM VPA for 8 hours. Chromatin was extracted from control and treated cells and immunoprecipitated with antibodies raised against three histone modifications associated with active transcription (H3K9ac, H4K8ac, H3K4me3) and one associated with repressed transcription (H3K27me3). DNA in bound (B) and unbound (UB) fractions were analysed by qPCR using region-specific primers (1.1 = *hoxb1* promoter; 1.2 = *hoxb1* transcription start site; 1.3 = *hoxb1* first exon; Similarly for *hoxb2*, *4*, *5* and *7*). The relative enrichment was calculated as the ratio B/UB and the mark was considered as present if this ratio was equal or above 1. Error bars represent the standard error of the mean (n=2).

3.3.7 Summary

These experiments were designed to examine the role of polycomb silencing in the transcriptional response to VPA treatment. Expression of genes in the *hoxb* cluster was assessed before and after VPA treatment in a wild-type ES cell line as well as in an ES cell line knocked out for the two polycomb proteins *Eed* and *Ring1b*. As expected, global levels of the silencing mark H3K27me3 decreased where expression of *hoxb* genes was up-regulated. Interestingly, loss of H3K27me3 was concomitent to a dramatic increase in H3K27ac. With the exception of *hoxb1* in dko cells, VPA treatment failed to further up-regulate *hoxb* genes, suggesting the limited effect of histone hyperacetylation on transcription. The transcriptional activation of hoxb1 could reflect early differentiation in dko cells rather than a direct effect of HDAC inhibition. In wt cells, VPA treatment induced the up-regulation of *hoxb1*, *4* and *7*. However, ChIP analysis of *hoxb* genes did not provide a clear link between transcriptional activities and histone modifications. H3K4me3, H3K27me3 and H4K8ac remained mostly unaffected by VPA treatment whereas H3K9ac was consistently up-regulated at all *hoxb* genes. Interestingly, *hoxb4* and *7* (up-regulated by VPA) were the only genes enriched for activating marks in untreated cells.

4. Discussion

4.1 Do histone modifications code for epigenetic information?

Epigenetic research has rapidly expanded in the last 50 years, and the high complexity of mechanisms described so far perhaps requires revisiting its fundamentals. The current consensus is that epigenetic mechanisms allow the establishment and the long-term maintenance of specific gene expression patterns. It was proposed that, without changing the DNA sequence, these mechanisms establish the expression of different programmes of genes in cells that carry the same genetic material. Thereby, one genome can produce various cell types in the development of a complex organism (Holliday, 2006).

At first, the observation of different levels of DNA compaction in the nucleus emphasised the idea of DNA sequence-independent mechanism(s) for gene regulation. In fact, the more the DNA is compacted, the less it is accessible to regulatory proteins like transcription factors. Electron microscopy and crystallography studies allowed the discovery of the chromatin fibre, a result of the interaction between the DNA molecule and histone proteins (Olins and Olins, 1974; Widom, 1986). Chromatin was revealed to be an organised structure when the nucleosome was characterised, eight histone proteins around which 146bp of DNA are wrapped (Luger et al., 1997). It was therefore suggested that the arrangement of nucleosomes would dictate degrees of chromatin compaction in different regions of the genome (Lee et al., 2007c). Importantly, many years of research showed the existence of cytosine methylation, post-translational modification of histones tails, incorporation of histone variants or RNAi-mediated regulation. All those events were classified as

epigenetic processes because of their potential short and long-term impact on gene transcription.

Unquestionably, specific histone modifications and DNA methylation have been correlated with transcriptional activity. In fact, maps of key histone modifications on active and inactive genes have been defined using high throughput technologies. For instance, actively transcribed genes usually show high levels of histone acetylation or H3K4 tri-methylation around their transcription start site (TSS) while H3K27me3 and H3K9me3 were mainly found in silent genes (Barski et al., 2007). Histone modifications were proposed to regulate transcription via direct or indirect mechanisms, either by altering histone/DNA interactions (Tse et al., 1998) or by providing binding sites for regulatory proteins (Taverna et al., 2007). Furthermore, histone marks were suspected to regulate transcription through coordinated action, thus constituting a “histone code” where a specific group of modifications alter transcription in a distinct way (Jenuwein and Allis, 2001). Importantly, the term “epigenetic” involves inheritance through the cell cycle. Persistent patterns of DNA methylation in imprinting or X-inactivation, maintenance of chromatin states after mitosis in lineage-committed cells and the existence of bivalent domains in stem cells provid good arguments for the inheritance of epigenetic marks and their role in the long-term maintenance of gene expression patterns (Kaufman and Rando, 2010).

In parallel to the development of the “nucleosome signalling” hypothesis, another literature has rapidly grown, pointing to the lack of evidence for a direct role for histone modifications in gene regulation. In fact, although specific marks have clearly been correlated with transcriptional activity, a causative role has not been proven. In addition, how histone modifications survive replication and mitosis remains largly

unclear. Therefore, it is only fair to wonder if histone modifications strictly constitute “epigenetic” information (Ptashne, 2007).

If histone modifications are a cause rather than a consequence of gene activity, and if they are responsible for the memory of transcription states, they should (1) be able to target regulatory proteins where there are needed, (2) be sufficient for activation or silencing and (3) be retained through mitosis, before regulatory proteins re-associate at loci (Ringrose and Paro, 2007).

In order to address some of those issues, we choose in this study to alter histone acetylation using an HDAC inhibitor and measure the short and long-term impact on global gene expression and levels of various histone modifications. Mouse ES cells were treated with valproic acid (VPA), an inhibitor of class I and IIa HDACs. Because VPA affects many HDACs, the treatment was expected to induce global hyperacetylation. Acetylation of histones has been widely associated with active transcription, therefore we hypothesised there would be important changes in gene expression following VPA treatment. In addition, if histone modifications were faithfully retained during the cell cycle, hyperacetylation and changes in gene expression should persist through to subsequent generations of cells, although VPA had been removed from the culture media. We thus decided to test this idea via washout experiments. Finally, we also turned our interest on the hox genes, VPA being a powerful teratogen (Nau et al., 1991). It was therefore interesting to assess the link between deregulation of hoxb genes and potential hyperacetylation. In fact, hox genes can be induced during ES cell differentiation and therefore offer a good model to study the potential predictive role of histone modifications.

4.2 Hyperacetylation in ES cells: Consequences on transcription and histone modifications

4.2.1 The lack of evidence for a causative role

High throughput experiments have highlighted a strong correlation between active transcription and acetylation of H3 and H4 (Roh et al., 2006). In addition, many studies using differentiation systems in vitro have shown that acetylation at transcriptional start sites correlates with cell- and stage-specific transcription as well as transcriptional poising (Cruickshank et al., 2010). However, as strong as the correlations can appear, they do not give evidence for a causative effect of acetylation.

One putative direct effect of acetylation is the neutralisation of histone tail lysine positive charge and thereby the alteration of DNA/histone interactions leading to the opening of the chromatin fibre. It was proposed that chromatin opening would increase DNA accessibility to proteins including transcription factors, co-factors or even RNA polymerase II (Lee et al., 1993; Vettese-Dadey et al., 1996). Recombinant acetylated nucleosomal arrays in fact show reduced condensation (Tse et al., 1998) and transcription was affected by charge neutralisation of specific lysine residues (Dion et al., 2005). However, very few studies have explored the relationship between histone acetylation and DNA accessibility. A classic method exists in the mapping of DNase I hypersensitive (DHP) sites using high throughput sequencing. Interestingly DHP sites were found in promoters, enhancers, insulators or locus control regions. Importantly, the presence of DHP sites correlated with RNA pol II binding and active gene expression (Boyle et al., 2008). More recently, the measure of DNA accessibility using a methylation foot printing assay in Drosophila was

performed. Once again, accessibility was higher at the TSS of active genes and coincided with peaks of RNA pol II, H3K4me3 and H4K16ac. A particular role of H4K16ac was highlighted as the mark strongly correlated with early replication and high accessibility. On the other hand, the lowest accessibility was found in regions flanked by H3K27me3 (Bell et al., 2010). It is worth noted that DHP sites were also found in silent genes, along with active histone marks and RNA pol II, possibly as marks of poised genes (Boyle et al., 2008). Those studies seem to confirm the role of acetylation in chromatin opening and DNA accessibility. Although it is very tempting to conclude that acetylation can increase DNA accessibility, those studies still did not prove the causative role of this mark.

The discovery of bromodomains, conserved domains capable of binding to acetylated residues, gave credit to the additional idea that histone modifications provide docking sites for regulatory proteins (Taverna et al., 2007). The human genome encodes for 42 bromodomain-containing proteins including co-factors, chromatin modifiers or DNA helicases (Sanchez and Zhou, 2009). For example, bromodomains were found in regulatory proteins such as TAFII250 (Jacobson et al., 2000) or the chromatin remodeler Swi2/Snf2 (Chandy et al., 2006). Once again, a simple model in which regulatory proteins bind to a specific mark in order to modulate transcription is easily conceivable. However, modifying enzymes rarely contain a DNA-binding domain and often need to be recruited by other proteins, which suggests a secondary role of acetylation rather than a decisive role. For instance, the histone acetyl-transferases CBP and p300 interact with at least 400 proteins, many of them being transcription factors (Bedford et al., 2010). Nevertheless, knockout of HATs in mice are generally lethal (Yamauchi et al., 2000; Yao et al., 1998).

Thus, although acetylation seems to be indispensable, it appears to act downstream of original signals for a change in transcription. To test this hypothesis, the first question we have addressed in this study is whether or not hyperacetylation on its own can affect gene activity.

4.2.2 The consequences of hyperacetylation on global gene expression

Western blotting confirmed that 1mM VPA treatment was enough to induce a massive and fast increase of global acetylation levels on histone H3, H4 and H2B. The potential cross-talk between acetylation and methylation in response to VPA, which was found in Human promyelocytic leukemia cells (Nightingale et al., 2007), was also observed in ES cells, as VPA treatment was able to raise global levels of H3K4me3. However, H3K27me3 was not affected, at least at a global level.

ES cells were treated for eight hours, maximising histone hyperacetylation. Microarray technology was then used to measure global gene expression. Stringent statistical analysis ($p<0.1$ after FDR correction) revealed that transcription of 1144 genes was affected by the treatment. This represents only 10% of the total number of genes on the array, which is significant but not as large an effect as expected considering the dramatic increase in acetylation induced by VPA. Furthermore, no more than 123 genes showed a change greater than 2-fold (95 up and 28 down), suggesting a very limited effect of hyperacetylation on transcription. Small changes following HDACi treatment have already been observed in the past (Dannenberg and Edenberg, 2006; Joseph et al., 2004).

Chromatin in ES cells is generally highly acetylated compared to lineage-committed cells (Meshorer et al., 2006), therefore it is possible that another level of hyperacetylation would not be effective for further transcriptional changes. However, studies in our lab showed similar effects in Lymphoblastoid cells following VPA treatment, suggesting that limited transcriptional change following VPA treatment is not a characteristic of undifferentiated cells. Another potential explanation for the limited effects resides in the fact that 95% of the genome in complex organism is not protein-coding. However, those non-coding sequences are not functionless as they contain important elements for gene regulation including non-coding RNA or retrotransposons (Taft et al., 2007). Hyperacetylation at these loci could therefore generate indirect effects. Also, in favour of the histone code hypothesis, where a mark may act in unison with others and thus have a limited effect on its own, hyperacetylation could be simply not enough to impact on gene activity.

Finally, in spite of the fact that acetylation has been associated with active transcription, the VPA treatment induced both up and down regulation of genes. This could first be because VPA interferes with several signalling pathways (Kostrouchova et al., 2007). More importantly, a wide range of proteins, including transcription factors, are substrates to the action of HDACs (Kim et al., 2006). In fact, butyrate treatment was shown to affect the binding of some transcription factors (Rada-Iglesias et al., 2007). Furthermore, it is difficult to distinguish the primary effects of histone hyperacetylation and secondary effects resulting from the modulation of acetylated proteins. These phenomena are likely to explain the large proportion of genes down-regulated by VPA.

4.2.3 Global changes versus local changes

Even though western blots revealed important increases in VPA-induced acetyl mark abundance, local patterns could vary from one gene to the other. In fact, it is rather unlikely that all genes are regulated by HDAC enzymes, so one could expect the hyperacetylation not to be uniform across the genome. Binding maps of five HATs (p300, CBP, MOF, PCAF, Tip60) and four HDACs (1, 2, 3, 6) in $CD4^+$ T cells were generated using ChIP-sequencing (Wang et al., 2009). Surprisingly, occupancy of all of these enzymes positively correlated with gene expression, RNA pol II binding and H3/H4 acetylation. Importantly, treatment with the HDACi trichostatin A (TSA) induced H3K9 and H4K16 hyperacetylation mostly at active genes or silent genes enriched in H3K4me3 and suspected to be poised for transcription. However, TSA increased Pol II binding at silent genes enriched for H3K4me3 but did not induce transcription (Wang et al., 2009). The limited effect of a VPA treatment on transcription in ES cells could therefore be due to (1) the fact that HDACs are only active on a subset of genes and (2) the fact that hyperacetylation alone is not sufficient for the activation of silent genes.

In addition, one could imagine several mechanisms of gene protection against hyperacetylation. For instance, important regulatory regions could be protected by insulators. Hyperacetylation could also be buffered by other modifications, or redundancy between HDACs could allow the maintenance of normal acetylation levels where it is important. In fact, VPA targets many HDACs but not all. Therefore, assessing the degree of mark enrichment at specific genes would be essential to understand their role in HDACi-mode of action. For this purpose, ChIP technology was used and allowed quantification of specific histone marks at specific regions of

the genome. Levels of the three “activating” marks, (H3K9ac, H4K8ac and H3K4me3), and the “silencing” mark (H3K27me3), were analysed across several genes that showed various responses to the VPA treatment: *Egr1*, *Ndgr4* and *H1f0* were up-regulated whereas *Anapc7* and *Slc6a8* were unaffected and *Nanog* was down-regulated.

Because we looked at a small number of genes, and qPCR only allows the analysis of small regions of the genome, it is very difficult to draw general conclusions from those experiments. However, we were able to detect trends within this specific pool of genes: First of all, levels of H3K9 acetylation were raised by VPA on all genes regardless of the transcriptional response to the treatment. Although levels of enrichment in untreated cells did not correlate with the fold changes after treatment, increases on non-affected genes and *Nanog* were modest compared with increases at up-regulated genes. More importantly however, while we showed by western blotting that H3K9ac was increased by 16-fold following VPA treatment, ChIP data never showed enrichment by more that 2.5-fold. These results suggest that the genes we analysed were protected against massive hyperacetylation. It is also possible that this set of genes was not targeted by HDACs to start with, so it would be interesting to measure the local abundance of these enzymes by ChIP.

Even though levels of H4K8ac were low in untreated cells, loss of enrichment was consistently observed and seemed to be proportional to the H3K9ac increase. This result is consistent with a previous study in which hepatocarcinoma cells were treated with the HDACi butyrate. The authors reported frequent H4 deacetylation

after the treatment (Rada-Iglesias et al., 2007). However, H3 also appeared deacetylated and the authors concluded that hyperacetylation was only an early and transient event (occurring between 2 and 4 hours of treatment). In this study, deacetylation of both H3 and H4, along with loss of RNA pol II binding, correlated with down-regulation of the genes. Interestingly, the authors showed loss of binding of several factors including CBP at specific genes. That could well explain the loss of acetylation following VPA treatment (Rada-Iglesias et al., 2007). In our study however, we did not detect H3 deacetylation, and deacetylation of H4 occurred even on up-regulated genes. As we looked at only a small number of genes, this may not reflect a general trend.

Histone acetylation and methylation have shown some degree of interdependence (Daniel et al., 2004; Dou et al., 2005) so it was interesting to measure the impact of a VPA treatment on some methyl marks. Working with ES cells, the enrichments of the "activating" mark H3K4me3 and the "silencing" mark H3K27me3 were measured. Although fold changes were not massive, it was interesting to see that H3K4me3 only increased at the promoter of the two up-regulated genes *Egr1* and *Ndrg4*. On the other hand, significant losses of enrichment were observed at the promoter of *Nanog* (down-regulated by the treatment) and *Anapc7* that was unaffected by VPA. It could therefore be speculated that the effect of VPA on transcription is translated by the histone mark H3K4me3 rather than hyperacetylation. A high throughput study would be essential to test this hypothesis.

Finally, consistently with western blot analysis, levels of H3K27me3 were low and were not affected by VPA treatment. Interestingly, the mark was more abundant at the promoters of up-regulated genes, particularly *Egr1* and *H1f0*. This might suggest a certain level of bivalency at these genes and that they were poised for transcription. In this case, VPA would potentially have an enhancer role. The analysis of bivalent genes like the Hox genes was thus a possible next step.

4.2.4 Hyperacetylation and pluripotency

Because this study partly aimed to measure the long term effect of transient VPA-induced hyperacetylation in ES cells, it was essential to limit "side" effects included differentiation. In fact, if the cells had differentiated after the treatment, potential long-term effects could be the result of the loss of pluripotency rather than the memory of hyperacetylation. In addition, HDAC inhibitors are believed to combat malignancy by inducing apoptosis, cell cycle arrest, differentiation (Kuendgen and Gattermann, 2007) or the reactivation of tumour suppressor genes (Takai and Narahara, 2008). With the concept of cancer-stem cells currently growing (Clevers, 2011), it was therefore interesting to analyse the global effect of VPA treatment on the ES cell transcriptome.

Gene clustering analysis showed that eight hours VPA treatment significantly affects the expression of genes involved in cell cycle regulation or apoptosis. However, cell cycle profiles of treated ES cells obtained by FACS analysis showed that VPA has a very limited effect. Futhermore, measure of cell death by trypan blue staining did not show any significant changes. These results suggest that the changes in

transcription were compensated later. It is also important to remember that microarray analysis only allows one to measure effects at a transcriptional level. In fact, variations in transcript abundance do not mean variations in protein abundance.

Interestingly, the expression of some histone modifying enzymes, including lysine demethylases and lysine acetyl-transferases, was down-regulated. This could explain the limited changes observed by ChIP at gene promoters. However, this phenomenon could not be global as western blotting experiments confirmed the massive hyperacetylation induced by VPA. We were however unsuccessful in detecting hyperacetylated loci. Interestingly, immunofluorescent staining showed that hyperacetylation of H3/H4 following butyrate treatment mostly occurs at the nuclear periphery, which suggests that hyperacetylation could take place on constitutive heterochromatin or cytoplasmic histones (Rada-Iglesias et al., 2007).

Finally, at a transcriptional level, very few genes involved in the regulation of pluripotency and differentiation were affected by VPA. However, transcription of Nanog, Klf4 and Sox2, master regulators of pluripotency, was modulated in response to the treatment. Nevertheless, VPA treatment did not affect OCT4 and KLF4 protein levels and the transcriptional effects were only transient.

Although the idea that HDACi affect transcription to inhibit the proliferation of cancer cells is attractive, it seems that cells were able to maintain their phenotype even when exposed to the drug. The analysis of complete transcriptome and proteome would be useful to draw final conclusions. However, our data confirmed that gene expression involves more than one level of regulation and systems must exist to ensure proper cell function in different conditions. For instance here, it seems that

hyperacetylation does not occur at important regulatory sequences (i.e. promoters). In addition, microarray data revealed that affected genes were both enhancers and inhibitors of functions like cell cycle progression or apoptosis, which could result in limited changes of the phenotype. In addition, changes may be needed to preserve the normal cell cycle under a different environment such as metabolic changes caused by VPA treatment. It is important to remember that these observations are only valid for eight hours treatment with 1mM VPA. Effects including cell cycle arrest or apoptosis could still be caused by longer treatment.

It is also possible that HDAC inhibition acts via the modulation of non-histone protein acetylation. Several studies have shown that acetylation is an abundant post-translational modification (Choudhary et al., 2009). For instance, a highly conserved lysine residue was found in the nuclear export signal of Sox2. Acetylation at this site is required for the nuclear export of Sox2 and its degradation via ubiquitination and proteasome degradation (Baltus et al., 2009b). The up-regulation of Sox2 we observed after VPA treatment could then be the result of an increase in Sox2 degradation. Interestingly, expression of Nanog was shown to be regulated by Sox2 that brings the mSin3A-HDAC complex to the Nanog promoter. Consistent with our data, loss of activity of the complex led to a decrease of Nanog expression (Baltus et al., 2009a).

4.3 Response of homeobox genes to a VPA treatment

4.3.1 Histone modifications and transcription

The Hox gene family, coding for homeobox transcription factors, plays a crucial role in the differentiation process (Krumlauf, 1994). Importantly in ES cells, these genes are maintained silent but hypothetically poised for activation via the establishment of bivalent domains at their promoters. The presence of H3K4me3 might predict rapid activation upon specific signals, whereas H3K27me3 is believed to ensure silencing (Bernstein et al., 2006). As VPA is a powerful teratogen (Nau et al., 1991), it was interesting to test if hyperacetylation could disrupt this bivalency and the expression state of Hox genes.

Although levels of Hoxb gene expression were very low in ES cells, we observed the up-regulation of Hoxb1, 4 and 7 following VPA treatment in two different ES cell lines. We therefore decided to analyse histone modifications at the promoter of hoxb1, 4 and 7 as well as hoxb2 and 5 that were not affected by the treatment. Levels of H3K9ac, H4K8ac, H3K4me3 and H3K27me3 were assessed by ChIP before and after eight hours treatment with VPA. Strikingly, hoxb4 and hoxb7 were the only genes showing bivalency around the transcription start site. Although levels of H3K27me3 were high at all studied hoxb genes before treatment, only hoxb4 and hoxb7 showed enrichment for H3K4me3. Those observations contradict the current consensus that describes hoxb genes as well-known bivalent genes. Bivalent genes have been described as genes enriched for both the activating mark H3K4me3 and the silencing mark H3K27me3. Co-localisation of these marks is believed to provide a mechanism for rapid activation or rapid silencing of a gene. However, because the

regulation of hox genes appear to depend on their clustering, it is possible that bivalency is not needed at each promoter. Interestingly, it appears that hoxb4 and 7 were also the only genes enriched for H3K9ac and to less extent H4K8ac.

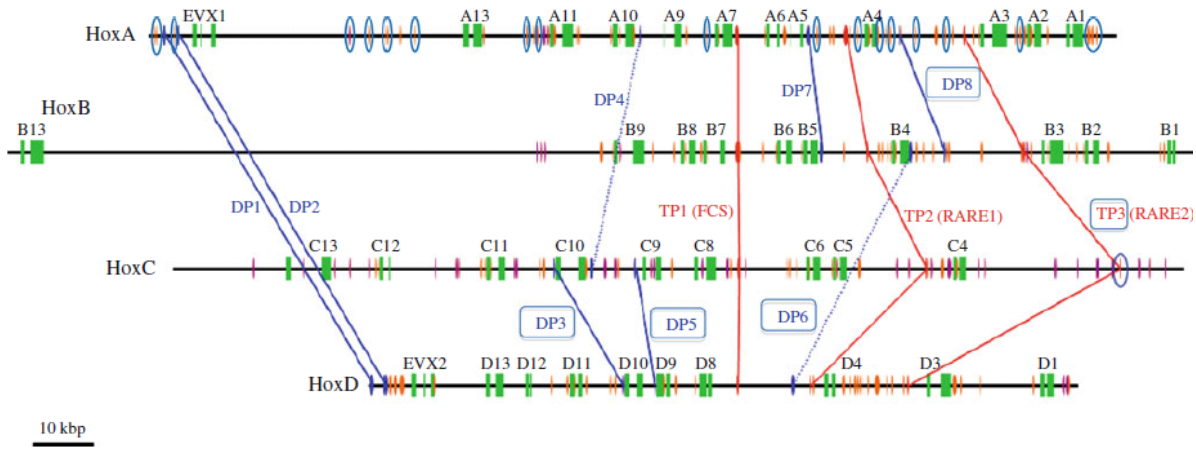
After treatment, H3K9ac was increased at all promoters with significant fold changes. On the other hand, H4K8ac and H3K4me3 levels were not affected by the treatment suggesting no role for these marks in the response to VPA at the hoxb cluster. H3K27me3 levels remained very high after treatment, although modest deceases were sporadically observed. With the exception of H3K9ac that significantly increased after treatment, regardless of the transcriptional state of genes, the histone marks we analysed remained essentially unchanged following VPA treatment. These results again strongly suggest that hyperacetylation does not drive transcription on its own. Interestingly, the only genes that appeared bivalent (i.e. hoxb4 and 7) were up-regulated by the treatment. We speculate that levels of histone modifications after treatment do not drive transcription, but levels at the start predispose genes for transcriptional change. However, hoxb1 was also up-regulated by VPA and showed very low levels of H3K9ac, H4K8ac, and H3K4me3. Analysis of additional regions could be informative as it is possible that cis-regulatory elements related to hoxb1 drove its transcriptional response to the treatment via the modulation of histone modifications. In fact, the comparison of hox clusters from 19 vertebrate species identified 208 conserved non-coding regions including 43 in the hoxb cluster (Matsunami et al., 2010). Phylogenetic footprinting was then used to identify binding sites in non-coding sequences. Many of these regions had previously been known to generate micro-RNA or to be cis-regulatory elements, however 160 new elements were detected including a new RARE motif potentially associated to hox4

(Matsunami et al., 2010) **(Figure 4.1)**. Acetylation at RARE has been shown to occur prior the induction of hox genes (Chambeyron and Bickmore, 2004; Gillespie and Gudas, 2007), therefore how VPA modulates histone modification at non-coding sequences in the hoxb cluster could be crucial for the understanding of transcriptional changes.

Another element that could have been responsible for the behaviour of hoxb genes before and after treatment is the spatial organisation of the cluster. Very little has been published on the three-dimensional (3D) architecture of hox clusters. A recent study attempted to characterise the 3D organisation of human hox clusters in embryonal carcinoma cells (Ferraiuolo et al., 2010). Using chromosome conformation capture (3C), the authors identified regions of elevated interaction frequency and thereby predicted looping of the clusters **(Figure 4.2)**. In the hoxA cluster, looping was more important in the 5' region and was lost by the addition of retinoic acid and to a lesser extend by EZH2 depletion. Finally in this study, the insulator CTCF was identified as a potential mediator of DNA loops (Ferraiuolo et al., 2010). Unfortunately, the equivalent study has not been done in mouse. However, recent ENCODE data are starting to be available and provide some information on the hox clusters. Regarding histone modifications, only H3K4me3 data are currently displayed. In agreement with our results, enrichments at the TSS of hoxb4 and 7 are the highest. Hoxb2 appears depleted at the TSS but enriched in the gene body, which we could not have detected with our primers. The insulator CTCF appears to bind more often at the end of the cluster, which correlates with what was observed in the Ferraiuolo et al. study. Interestingly, hoxb1, 4 and 7 are found in a region widely

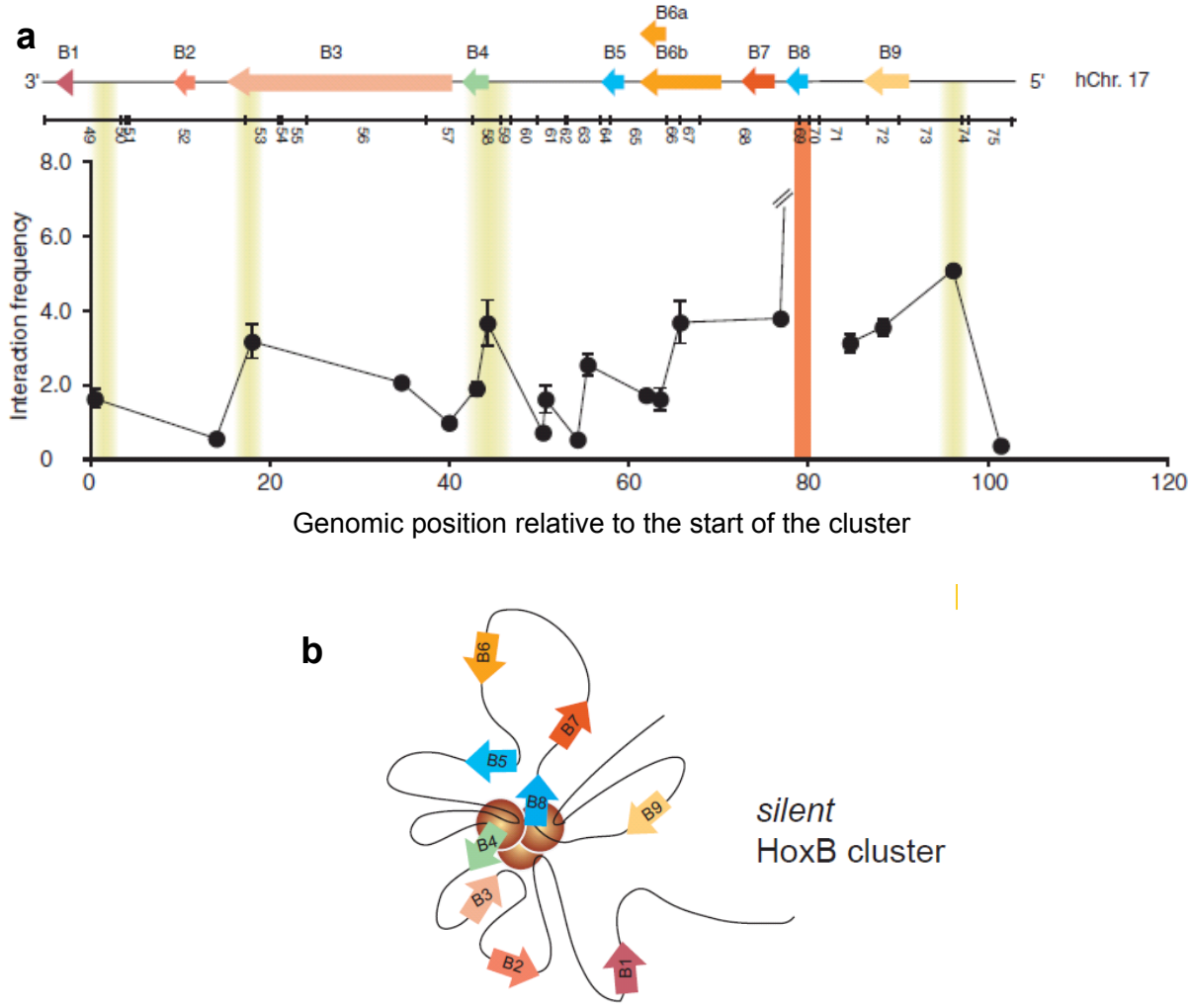
depleted in CTCF, which could partly explain their special behaviour. Finally, Pol II binding appears very low which correlates with silencing in ES cells **(Figure 4.3)**. High-resolution analysis of the cluster 3D-organisation before and after treatment would be highly informative for a potential role of looping and CTCF binding in the transcriptional response to VPA.

Figure 4.1 Conserved non-coding sequences within the four human hox cluters



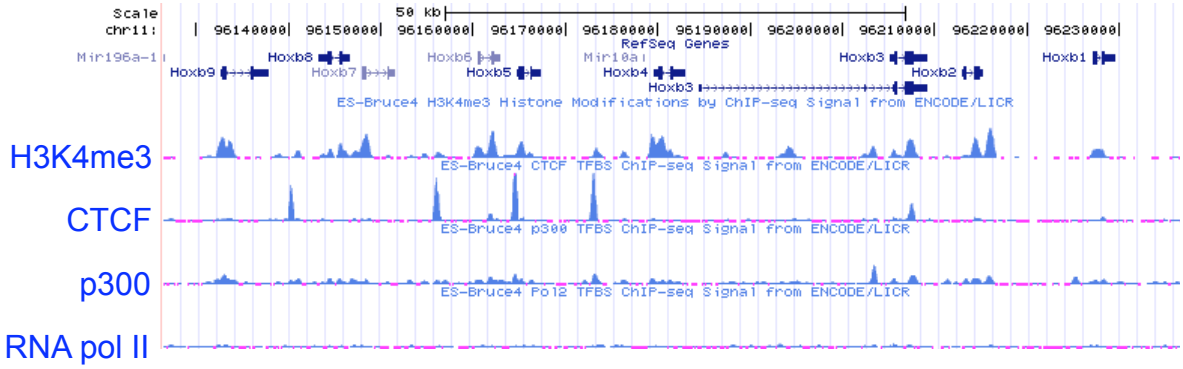
Conserved non-coding sequences (CNSs) were identified using sequence alignment of 19 vertebrate species. Here is represented the schematic diagram of CNSs over the four human hox cluters. 208 CNSs were detected including 160 new. Many of these sites constitute cis-regulatory elements or generate micro-RNA. Interestingly, two of the three tetra-paralogous CNS contain RARE motifs.
Exons of protein coding genes (Green boxes); Orthologous CNSs (Orange ovals); Tetra-paralogous (TP) and di-paralogous (DP) CNSs (Red and blue ovals); Newly detected CNSs (Light blue rectangles and circles).
Taken from Matsunami et al., 2010

Figure 4.2 Three dimensional representation of the human hoxb cluster



a, 3C interaction profile indicating looping contacts in the silent hoxb cluster. The solid orange vertical line represents the position of "fixed" 3C region.
b, Predicted two-dimensional representation of the hoxb cluster.
Taken from Ferraiuolo et al., 2010

Figure 4.3 Mouse hoxb cluster ENCODE data



The ENCODE (Encyclopaedia of DNA elements) project maps functional elements all over the human and recently mouse genome. Here are represented ChIP-seq data obtained for the Hoxb cluster including CTCF occupancy along with enrichment profiles of the active mark H3K4me3, the acetyl-transferase p300 and the RNA pol II in mouse embryonic stem cells.

4.3.2 Role of the Polycomb proteins

Results from our ChIP experiment confirmed that a “blanket” of H3K27me3 covers the hoxb cluster and is potentially responsible for the silencing of hoxb genes in ES cells. For many of the genes we studied, when hyperacetylation occurred at the promoter following VPA treatment, it was not sufficient to induce transcription. Therefore, we hypothesised that one protective mechanism involves the polycomb mark, which provided a silencing mechanism that hyperacetylation could not reverse. For this reason, the transcriptional states of hoxb genes in ES cells knocked out for the two polycomb proteins eed and ring1b were analysed before and after VPA treatment.

Depletion of polycomb proteins induced a decrease of H3K27me3 global level and a striking increase in global H3K27ac abundance. These results are in agreement with previous studies that identified H3K27 methylation and acetylation as exclusive marks. In *Drosophila* knockout of e(z), a switch between the two marks was observed globally and at the promoter of hox genes including *Abd-A* (Tie et al., 2009). Similar observations were made at the promoter of polycomb-target genes in mouse ES cells knocked out for Suz12 (Pasini et al., 2010). In both studies, CBP and trithorax proteins appeared necessary for the increase in H3K27ac. These results suggested that one role of the polycomb complex is to prevent H3K27 acetylation (Pasini et al., 2010). It would therefore be interesting to analyse levels of H3K27ac across the hoxb cluster, for instance at promoters or enhancers like RARE regions. Although VPA did not affect H3K27me3, western blot analysis of wild type and double knock-out cells

showed that global levels of H3K27ac were increased following treatment. Therefore, this mark could play an important role in the hox genes response to VPA.

In knockout cells, hoxb transcripts were generally increased. However, this induction was small compared with that observed during retinoic acid differentiation. Here again, the modulation of histone modifications was not sufficient to induce maximum transcription. Even though the silencing mark H3K27me3 was decreased and acetylation was enhanced, expression levels remained very low. We then tested if further hyperacetylation induced by VPA could derepress hoxb gene transcription. With the exception of hoxb1, that was significantly up-regulated, none of the hoxb genes were affected by the treatment. Unfortunately, we have not performed ChIP analysis of the cluster in these cell lines so we can not claim that acetylation was raised specifically at the hox genes following treatment in dko cells. If so, it would indicate again that acetylation is not a causative signal for transcription. However, it is possible that mechanism prevented further hyperacetylation at these sites. Finally, the increase in hoxb1 transcript could be due to early differentiation. It would be interesting to analyse histone marks at its promoter or at the associated RARE-enhancer. In fact, H3K27ac was shown to specifically occupy active enhancers (Creyghton et al., 2010).

4.4 Long term effect of transient hyperacetylation

4.4.1 Role of histone modifications in epigenetic inheritance

Histone modifications are believed to be heritable through the cell cycle for the maintenance of chromatin states and gene expression patterns. However, mechanisms of inheritance are not well understood. It has been hypothesised that histone modifications are put in place during replication, through interactions between modifying enzymes and the replication machinery (Probst et al., 2009). The role of histone chaperones in preventing the dilution of enzymes or histone proteins has been highlighted as well (Avvakumov et al., 2011; Quivy et al., 2008). In addition, past studies have shown that several binding proteins or histone modifications colocalise with metaphase chromosomes and therefore are potentially transmitted to daughter cells (Aoto et al., 2008; Blobel et al., 2009; Hansen et al., 2008; Terrenoire et al., 2010).

One way of testing the inheritance of histone modifications is to induce them with an environmental signal and assess if they can persist, even in the absence of the original signal. Very little evidence of the inheritance of induced marks is available to date, leading to some controversy concerning the epigenetic nature of histone marks. An early study in yeast showed that TSA treatment was able to induce derepression of centromeric genes along with chromosome segregation defects and hyperacetylation. Those effects persisted for several rounds of cell division after TSA removal, suggesting that hyperacetylation could be remembered (Ekwall et al., 1997). However, both hyperacetylation and transcription were persistent, introducing uncertainty over which is responsible for the other.

Another study claimed that relocation at the nuclear periphery was responsible for the transcriptional memory of Gal genes. The authors suggested that the histone variant H2A.Z was required (Brickner et al., 2007). However, experiments using heterokaryon cells – with two nuclei in the same cytoplasmic environment – revealed that the memory of previous transcription was dependent on a cytoplasmic element, Gal1-galactokinase (Zacharioudakis et al., 2007).

4.4.2 Memory in ES cells

Because so few studies have asked whether histone modifications could be the carrier of an epigenetic memory, we decided to test this hypothesis in our ES cell line. Cells were treated for eight hours with VPA and washed out for five days. Acetylation levels and global transcription were analysed after treatment and after washout in order to see if cells could remember hyperacetylation and induced changes in gene expression. HDAC enzymes have been shown to interact with PCNA suggesting that acetyl marks are deposited on incorporated histones during replication (Probst et al., 2009). In addition, active genes were associated with early replication and acetylation. Importantly, when gene templates were injected into early replicating nuclei, they became acetylated and transcribed and this state was preserved through the cell cycle (Zhang et al., 2002). It therefore appears that histone acetylation participates to the inheritance of chromatin states and transcription levels.

When gene expression in washed out and control cells were compared, we could not detect any significantly up- or down-regulated genes. After five days post wash out,

expression of all VPA-affected genes was back to normal. We concluded that the HDACi VPA could not induce long-term effects on transcription in ES cells, at least for the genes we analysed and over the time of treatment we applied. In fact, the time of treatment was perhaps too short to induce long-term effects. The turnover of acetyl mark is generally considered high and increases with acetylation levels which suggests that this modification is not very stable (Zee et al., 2010). Therefore, acetylation is possibly not an ideal mark for memory. This would partly explain the rapid disappearance of the mark and the complete recovery of normal transcription activity after wash out. Although histone demethylases have been characterised, methylation appears as a more stable mark with little turnover and could thus provide a mechanism for histone mark inheritance (Rice and Allis, 2001). However, although we did observe an increase of H3K4me3 following treatment at a global level, the mark rapidly disappeared during the wash out.

Finally, ES cells divide very rapidly which was an advantage for our model system. However, eight hours of treatment might have created heterogeneity of populations. Because cells were not synchronised for the experiment, it is possible that they didn't respond uniformly to the short VPA treatment. Even if some of the cells kept the memory of hyperacetylation, they were rapidly diluted after wash out.

4.4.3 Memory in differentiating cells

Using retinoic acid, it is possible to progressively induce the hox genes during ES cell differentiation. This group of genes therefore provide a good model for the study of predictive marks, as if histone modifications form an epigenetic code, they must be predictive of transcription states. In ES cells, it was shown that H3K9ac and

H3K4me3 enrichments were increased by retinoic acid treatment in the promoter of hoxb9 days before its transcriptional induction (Chambeyron and Bickmore, 2004). In addition, several histone modifications have been shown to be predictive of transcription in differentiating B-cells. In fact, a region in the λ5-VpreB1 locus was already acetylated on H3 and di-methylated on H3K4 in ES cells. At a later stage transcription factors including TAF10, TBP, Brg1 or PU.1 and the RNA pol II would bind the region. Finally in pre-B cells, histone marks and transcription factor binding spread and λ5 and VpreB1 genes became expressed (Szutorisz et al., 2005). Finally, a previous study in our laboratory has shown that hyperacetylation, induced by VPA in differentiating mouse embryos, persisted after wash out of the drug. Embryos were treated between the 8-cells stage and the morula stage. An increase in H4 acetylation and H3K4me3 was observed at the promoters of hoxb1 and hoxb9 and remained until the blastocyst stage (VerMilyea et al., 2009).

Our experiments in ES cells showed that VPA treatment was not able to induce long-term hyperacetylation or changes in gene expression. We concluded that acetylation on its own was not a sufficient signal for substantial transcriptional changes or epigenetic memory. However, because histone modifications seamed to be predictive of future transcriptional activation in ES cells, it was decided to test the effect of hyperacetylation added to the differentiation signal. Experiments in mouse embryos can detect persistent changes at the promoter of developmental genes. We therefore decided to use retinoic acid-differentiation of ES cells as a model and analyse the long-term effect of transient exposure to VPA on hoxb gene expression. Cells were treated for 16 hours with VPA (8 hours before and after LIF removal) and

retinoic acid was added two days later. Hoxb gene expression was quantified every day until day 8 of differentiation. The experiment was repeated three times and when results were merged we could not observe any differences in the timing of hoxb gene expression in control and treated cells. The current model is that retinoic acid (RA) binds to retinoic acid response elements (RARE) and leads to the loss of polycomb proteins binding and a consequential loss of H3K27me3. RA is also able to recruit acetyl-transferases including p300 and RNA pol II (Gillespie and Gudas, 2007; Lee et al., 2007a). An increase in H3/H4 acetylation is therefore an important step in the transcriptional activation of RA target genes during differentiation. Thus, it is possible that an overlap between the RA and VPA treatments might have been more efficient to induce long-term effects. In fact, removal of LIF along with the addition of TSA was shown to induce the hoxa1 gene but only at half the level of RA-induction (Lee et al., 2007a). In our model system, the end of the VPA treatment and the beginning of the RA treatment were separated by 40 hours, which considering our previous results, is likely to be sufficient to reset histone marks and gene expression. In this case, normal timing of hoxb gene expression is not surprising.

However, when separate replicates were analysed, two out of three did show early or greater induction of some hoxb genes. We could thus speculate that the timing of hoxb gene induction was disturbed by early hyperacetylation of promoters or associated RARE. In order to confirm this scenario, the experiment should be repeated and ChIP experiment performed to analyse the evolution of histone modifications at regulatory sequence during the differentiation time course. In fact, variations between time courses could also be due to very low expression of hox genes in ES cells, which could cause leaking expression. Finally, variations could be

the result of asynchronous population of cells differentiating or cells going down different differentiation pathways.

4.5 Conclusions

Although our study requires additional experiments in order to finalise some important results, it seems likely that histone acetylation is not a straightforward mechanism. In fact, hyperacetylation does not always correlate with transcriptional activation and is certainly not enough on its own to relieve polycomb silencing. The set of experiments we performed could lead to the conclusion that histone acetylation is not a causative event. However, it is important to remember that we have chosen to work with the HDACi VPA, which is not a specific inhibitor and induces the hyperacetylation of many non-histone proteins. It is therefore possible that its mechanism of action remains too complicated to be properly dissected. It is also important to note that VPA targets HDAC enzymes so the hyperacetylation could be classified as an indirect effect. Acetylation by acetyl-transferases could partly cause transcriptional activation whereas HDAC enzymes would for instance be involved in balancing acetylation levels and preventing over-production of transcripts. In this case, the removal of the drug is very likely to reset basal transcription activities. Another explanation for the absence of long-term effects resides in the fact that, although HDACs associate with the replication machinery, it is conceivable that they act in coordination with other factors or marks. Chromatin and transcriptional states have to survive replication to ensure cell identity. However, the replication phase provides a time window for changes. One could imagine that the presence or the absence of specific factors would dictate if an acetyl mark were to be transmitted to

the next cell generation. In this case, the removal of VPA would not be a sufficient signal for memory.

In conclusion, histone acetylation could provide an important but not sufficient signal for active transcription. Although it would not be the first signal or the decisive signal for activation, it could still be important for the maintenance of open chromatin and active transcription. The fact that a biological event needs several layers of control makes sense, as cells have to buffer their environment in order to function properly.

5. References

Ahmad, K. and Henikoff, S. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* **9**, 1191-200.

Alison, M. R., Lim, S. M. and Nicholson, L. J. (2011). Cancer stem cells: problems for therapy? *J Pathol* **223**, 147-61.

Amati, B., Alevizopoulos, K. and Vlach, J. (1998). Myc and the cell cycle. *Front Biosci* **3**, d250-68.

Aoto, T., Saitoh, N., Sakamoto, Y., Watanabe, S. and Nakao, M. (2008). Polycomb group protein-associated chromatin is reproduced in post-mitotic G1 phase and is required for S phase progression. *J Biol Chem* **283**, 18905-15.

Arya, G., Maitra, A. and Grigoryev, S. A. (2010). A structural perspective on the where, how, why, and what of nucleosome positioning. *J Biomol Struct Dyn* **27**, 803-20.

Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N. and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* **17**, 126-40.

Avvakumov, N., Nourani, A. and Cote, J. (2011). Histone chaperones: modulators of chromatin marks. *Mol Cell* **41**, 502-14.

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

Baltus, G. A., Kowalski, M. P., Tutter, A. V. and Kadam, S. (2009a). A positive regulatory role for the mSin3A-HDAC complex in pluripotency through Nanog and Sox2. *J Biol Chem* **284**, 6998-7006.

Baltus, G. A., Kowalski, M. P., Zhai, H., Tutter, A. V., Quinn, D., Wall, D. and Kadam, S. (2009b). Acetylation of sox2 induces its nuclear export in embryonic stem cells. *Stem Cells* **27**, 2175-84.

Banman, S. L., McFie, P. J., Wilson, H. L. and Roesler, W. J. (2010). Nuclear redistribution of TCERG1 is required for its ability to inhibit the transcriptional and anti-proliferative activities of C/EBPalpha. *J Cell Biochem* **109**, 140-51.

Bannister, A. J. and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res* **21**, 381-95.

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C. and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120-4.

Bantignies, F., Grimaud, C., Lavrov, S., Gabut, M. and Cavalli, G. (2003). Inheritance of Polycomb-dependent chromosomal interactions in Drosophila. *Genes Dev* **17**, 2406-20.

Barakat, T. S., Gunhanlar, N., Gontan Pardo, C., Achame, E. M., Ghazvini, M., Boers, R., Kenter, A., Rentmeester, E., Grootegoed, J. A. and Gribnau, J. (2010). RNF12 Activates Xist and Is Essential for X Chromosome Inactivation. *PLoS Genet* **7**, e1002001.

Barber, B. A. and Rastegar, M. (2010). Epigenetic control of Hox genes during neurogenesis, development, and disease. *Ann Anat* **192**, 261-74.

Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-37.

Bedford, D. C., Kasper, L. H., Fukuyama, T. and Brindle, P. K. (2010). Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. *Epigenetics* **5**, 9-15.
Bedford, M. T. (2007). Arginine methylation at a glance. *J Cell Sci* **120**, 4243-6.
Beisel, C. and Paro, R. (2011). Silencing chromatin: comparing modes and mechanisms. *Nat Rev Genet* **12**, 123-35.
Bell, O., Schwaiger, M., Oakeley, E. J., Lienert, F., Beisel, C., Stadler, M. B. and Schubeler, D. (2010). Accessibility of the Drosophila genome discriminates PcG repression, H4K16 acetylation and replication timing. *Nat Struct Mol Biol* **17**, 894-900.
Belmont, A. S. (2006). Mitotic chromosome structure and condensation. *Curr Opin Cell Biol* **18**, 632-8.
Bernatavichute, Y. V., Zhang, X., Cokus, S., Pellegrini, M. and Jacobsen, S. E. (2008). Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in Arabidopsis thaliana. *PLoS One* **3**, e3156.
Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K. et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-26.
Berstine, E. G., Hooper, M. L., Grandchamp, S. and Ephrussi, B. (1973). Alkaline phosphatase activity in mouse teratoma. *Proc Natl Acad Sci U S A* **70**, 3899-903.
Bialik, S. and Kimchi, A. (2009). Lethal weapons: DAP-kinase, autophagy and cell death: DAP-kinase regulates autophagy. *Curr Opin Cell Biol* **22**, 199-205.
Bielecka, A. M. and Obuchowicz, E. (2008). Antiapoptotic action of lithium and valproate. *Pharmacol Rep* **60**, 771-82.
Bilodeau, S., Kagey, M. H., Frampton, G. M., Rahl, P. B. and Young, R. A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev* **23**, 2484-9.
Bird, A. and Tweedie, S. (1995). Transcriptional noise and the evolution of gene number. *Philos Trans R Soc Lond B Biol Sci* **349**, 249-53.
Blobel, G. A., Kadauke, S., Wang, E., Lau, A. W., Zuber, J., Chou, M. M. and Vakoc, C. R. (2009). A reconfigured pattern of MLL occupancy within mitotic chromatin promotes rapid transcriptional reactivation following mitotic exit. *Mol Cell* **36**, 970-83.
Blomen, V. A. and Boonstra, J. (2011). Stable transmission of reversible modifications: maintenance of epigenetic information through the cell cycle. *Cell Mol Life Sci* **68**, 27-44.
Bolden, J. E., Peart, M. J. and Johnstone, R. W. (2006). Anticancer activities of histone deacetylase inhibitors. *Nat Rev Drug Discov* **5**, 769-84.
Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G. et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-56.
Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K. et al. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349-53.
Boyle, A. P., Davis, S., Shulha, H. P., Meltzer, P., Margulies, E. H., Weng, Z., Furey, T. S. and Crawford, G. E. (2008). High-resolution mapping and characterization of open chromatin across the genome. *Cell* **132**, 311-22.
Bradbury, C. A., Khanim, F. L., Hayden, R., Bunce, C. M., White, D. A., Drayson, M. T., Craddock, C. and Turner, B. M. (2005). Histone deacetylases in acute myeloid

leukaemia show a distinctive pattern of expression that changes selectively in response to deacetylase inhibitors. *Leukemia* **19**, 1751-9.
Brickner, D. G., Cajigas, I., Fondufe-Mittendorf, Y., Ahmed, S., Lee, P. C., Widom, J. and Brickner, J. H. (2007). H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. *PLoS Biol* **5**, e81.
Campanero, M. R., Armstrong, M. I. and Flemington, E. K. (2000). CpG methylation as a mechanism for the regulation of E2F activity. *Proc Natl Acad Sci U S A* **97**, 6481-6.
Campos, E. I., Fillingham, J., Li, G., Zheng, H., Voigt, P., Kuo, W. H., Seepany, H., Gao, Z., Day, L. A., Greenblatt, J. F. et al. (2010). The program for processing newly synthesized histones H3.1 and H4. *Nat Struct Mol Biol* **17**, 1343-51.
Capecchi, M. R. (2005). Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat Rev Genet* **6**, 507-12.
Caretti, G., Di Padova, M., Micales, B., Lyons, G. E. and Sartorelli, V. (2004). The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. *Genes Dev* **18**, 2627-38.
Cartwright, P., McLean, C., Sheppard, A., Rivett, D., Jones, K. and Dalton, S. (2005). LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development* **132**, 885-96.
Chambeyron, S. and Bickmore, W. A. (2004). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev* **18**, 1119-30.
Chandy, M., Gutierrez, J. L., Prochasson, P. and Workman, J. L. (2006). SWI/SNF displaces SAGA-acetylated nucleosomes. *Eukaryot Cell* **5**, 1738-47.
Chang, B., Chen, Y., Zhao, Y. and Bruick, R. K. (2007). JMJD6 is a histone arginine demethylase. *Science* **318**, 444-7.
Chateauvieux, S., Morceau, F., Dicato, M. and Diederich, M. (2010). Molecular and therapeutic potential and toxicity of valproic acid. *J Biomed Biotechnol* **2010**.
Cheung, P., Tanner, K. G., Cheung, W. L., Sassone-Corsi, P., Denu, J. M. and Allis, C. D. (2000). Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. *Mol Cell* **5**, 905-15.
Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V. and Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834-40.
Chuang, L. S., Ian, H. I., Koh, T. W., Ng, H. H., Xu, G. and Li, B. F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* **277**, 1996-2000.
Clapier, C. R. and Cairns, B. R. (2009). The biology of chromatin remodeling complexes. *Annu Rev Biochem* **78**, 273-304.
Clevers, H. (2011). The cancer stem cell: premises, promises and challenges. *Nat Med* **17**, 313-9.
Coqueret, O. (2002). Linking cyclins to transcriptional control. *Gene* **299**, 35-55.
Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A. et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* **107**, 21931-6.
Creyghton, M. P., Markoulaki, S., Levine, S. S., Hanna, J., Lodato, M. A., Sha, K., Young, R. A., Jaenisch, R. and Boyer, L. A. (2008). H2AZ is enriched at polycomb

complex target genes in ES cells and is necessary for lineage commitment. *Cell* **135**, 649-61.
Cross, S. H. and Bird, A. P. (1995). CpG islands and genes. *Curr Opin Genet Dev* **5**, 309-14.
Cruickshank, M. N., Besant, P. and Ulgiati, D. (2010). The impact of histone post-translational modifications on developmental gene regulation. *Amino Acids* **39**, 1087-105.
Cuthbert, G. L., Daujat, S., Snowden, A. W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P. D., Tempst, P., Bannister, A. J. et al. (2004). Histone deimination antagonizes arginine methylation. *Cell* **118**, 545-53.
Daftary, G. S. and Taylor, H. S. (2006). Endocrine regulation of HOX genes. *Endocr Rev* **27**, 331-55.
Daniel, J. A., Torok, M. S., Sun, Z. W., Schieltz, D., Allis, C. D., Yates, J. R., 3rd and Grant, P. A. (2004). Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J Biol Chem* **279**, 1867-71.
Dannenberg, L. O. and Edenberg, H. J. (2006). Epigenetics of gene expression in human hepatoma cells: expression profiling the response to inhibition of DNA methylation and histone deacetylation. *BMC Genomics* **7**, 181.
De Santa, F., Totaro, M. G., Prosperini, E., Notarbartolo, S., Testa, G. and Natoli, G. (2007). The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell* **130**, 1083-94.
Dekker, F. J. and Haisma, H. J. (2009). Histone acetyl transferases as emerging drug targets. *Drug Discov Today* **14**, 942-8.
Dikic, I., Wakatsuki, S. and Walters, K. J. (2009). Ubiquitin-binding domains - from structures to functions. *Nat Rev Mol Cell Biol* **10**, 659-71.
Dillon, S. C., Zhang, X., Trievel, R. C. and Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol* **6**, 227.
Dimitri, P., Caizzi, R., Giordano, E., Carmela Accardo, M., Lattanzi, G. and Biamonti, G. (2009). Constitutive heterochromatin: a surprising variety of expressed sequences. *Chromosoma* **118**, 419-35.
Dimitri, P., Corradini, N., Rossi, F. and Verni, F. (2005). The paradox of functional heterochromatin. *Bioessays* **27**, 29-41.
Dion, M. F., Altschuler, S. J., Wu, L. F. and Rando, O. J. (2005). Genomic characterization reveals a simple histone H4 acetylation code. *Proc Natl Acad Sci U S A* **102**, 5501-6.
Dong, K. B., Maksakova, I. A., Mohn, F., Leung, D., Appanah, R., Lee, S., Yang, H. W., Lam, L. L., Mager, D. L., Schubeler, D. et al. (2008). DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO J* **27**, 2691-701.
Dorigo, B., Schalch, T., Bystricky, K. and Richmond, T. J. (2003). Chromatin fiber folding: requirement for the histone H4 N-terminal tail. *J Mol Biol* **327**, 85-96.
Dou, Y., Milne, T. A., Tackett, A. J., Smith, E. R., Fukuda, A., Wysocka, J., Allis, C. D., Chait, B. T., Hess, J. L. and Roeder, R. G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* **121**, 873-85.

Du, J., Chen, T., Zou, X., Xiong, B. and Lu, G. (2010). Dppa2 knockdown-induced differentiation and repressed proliferation of mouse embryonic stem cells. *J Biochem* **147**, 265-71.
Duester, G. (2008). Retinoic acid synthesis and signaling during early organogenesis. *Cell* **134**, 921-31.
Ekwall, K., Olsson, T., Turner, B. M., Cranston, G. and Allshire, R. C. (1997). Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres. *Cell* **91**, 1021-32.
English, C. M., Adkins, M. W., Carson, J. J., Churchill, M. E. and Tyler, J. K. (2006). Structural basis for the histone chaperone activity of Asf1. *Cell* **127**, 495-508.
English, C. M., Maluf, N. K., Tripet, B., Churchill, M. E. and Tyler, J. K. (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**, 13673-82.
Erfurth, F. E., Popovic, R., Grembecka, J., Cierpicki, T., Theisler, C., Xia, Z. B., Stuart, T., Diaz, M. O., Bushweller, J. H. and Zeleznik-Le, N. J. (2008). MLL protects CpG clusters from methylation within the Hoxa9 gene, maintaining transcript expression. *Proc Natl Acad Sci U S A* **105**, 7517-22.
Evans, M. J. and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-6.
Fan, Y., Nikitina, T., Morin-Kensicki, E. M., Zhao, J., Magnuson, T. R., Woodcock, C. L. and Skoultchi, A. I. (2003). H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo. *Mol Cell Biol* **23**, 4559-72.
Ferraiuolo, M. A., Rousseau, M., Miyamoto, C., Shenker, S., Wang, X. Q., Nadler, M., Blanchette, M. and Dostie, J. (2010). The three-dimensional architecture of Hox cluster silencing. *Nucleic Acids Res* **38**, 7472-84.
Fischle, W., Tseng, B. S., Dormann, H. L., Ueberheide, B. M., Garcia, B. A., Shabanowitz, J., Hunt, D. F., Funabiki, H. and Allis, C. D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438**, 1116-22.
Fischle, W., Wang, Y. and Allis, C. D. (2003). Binary switches and modification cassettes in histone biology and beyond. *Nature* **425**, 475-9.
Francis, N. J., Follmer, N. E., Simon, M. D., Aghia, G. and Butler, J. D. (2009). Polycomb proteins remain bound to chromatin and DNA during DNA replication in vitro. *Cell* **137**, 110-22.
Francis, N. J., Kingston, R. E. and Woodcock, C. L. (2004). Chromatin compaction by a polycomb group protein complex. *Science* **306**, 1574-7.
Frischer, L. E., Hagen, F. S. and Garber, R. L. (1986). An inversion that disrupts the Antennapedia gene causes abnormal structure and localization of RNAs. *Cell* **47**, 1017-23.
Frolov, M. V. and Dyson, N. J. (2004). Molecular mechanisms of E2F-dependent activation and pRB-mediated repression. *J Cell Sci* **117**, 2173-81.
Fu, Y., Sinha, M., Peterson, C. L. and Weng, Z. (2008). The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet* **4**, e1000138.
Gillespie, R. F. and Gudas, L. J. (2007). Retinoid regulated association of transcriptional co-regulators and the polycomb group protein SUZ12 with the retinoic acid response

elements of Hoxa1, RARbeta(2), and Cyp26A1 in F9 embryonal carcinoma cells. *J Mol Biol* **372**, 298-316.
Gottlicher, M., Minucci, S., Zhu, P., Kramer, O. H., Schimpf, A., Giavara, S., Sleeman, J. P., Lo Coco, F., Nervi, C., Pelicci, P. G. et al. (2001). Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells. *EMBO J* **20**, 6969-78.
Gould, A., Morrison, A., Sproat, G., White, R. A. and Krumlauf, R. (1997). Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping Hox expression patterns. *Genes Dev* **11**, 900-13.
Groth, A., Corpet, A., Cook, A. J., Roche, D., Bartek, J., Lukas, J. and Almouzni, G. (2007). Regulation of replication fork progression through histone supply and demand. *Science* **318**, 1928-31.
Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. and Young, R. A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77-88.
Haberland, M., Montgomery, R. L. and Olson, E. N. (2009). The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nat Rev Genet* **10**, 32-42.
Hansen, J. C., Ausio, J., Stanik, V. H. and van Holde, K. E. (1989). Homogeneous reconstituted oligonucleosomes, evidence for salt-dependent folding in the absence of histone H1. *Biochemistry* **28**, 9129-36.
Hansen, K. H., Bracken, A. P., Pasini, D., Dietrich, N., Gehani, S. S., Monrad, A., Rappsilber, J., Lerdrup, M. and Helin, K. (2008). A model for transmission of the H3K27me3 epigenetic mark. *Nat Cell Biol* **10**, 1291-300.
Heintzman, N. D., Hon, G. C., Hawkins, R. D., Kheradpour, P., Stark, A., Harp, L. F., Ye, Z., Lee, L. K., Stuart, R. K., Ching, C. W. et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108-12.
Henikoff, S., Furuyama, T. and Ahmad, K. (2004). Histone variants, nucleosome assembly and epigenetic inheritance. *Trends Genet* **20**, 320-6.
Hermann, A., Goyal, R. and Jeltsch, A. (2004). The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem* **279**, 48350-9.
Holliday, R. (2006). Epigenetics: a historical overview. *Epigenetics* **1**, 76-80.
Honore, B., Rasmussen, H. H., Vorum, H., Dejgaard, K., Liu, X., Gromov, P., Madsen, P., Gesser, B., Tommerup, N. and Celis, J. E. (1995). Heterogeneous nuclear ribonucleoproteins H, H', and F are members of a ubiquitously expressed subfamily of related but distinct proteins encoded by genes mapping to different chromosomes. *J Biol Chem* **270**, 28780-9.
Hu, X. V., Rodrigues, T. M., Tao, H., Baker, R. K., Miraglia, L., Orth, A. P., Lyons, G. E., Schultz, P. G. and Wu, X. (2010). Identification of RING finger protein 4 (RNF4) as a modulator of DNA demethylation through a functional genomics screen. *Proc Natl Acad Sci U S A* **107**, 15087-92.
Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W. and Melton, D. A. (2008). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotechnol* **26**, 1269-75.
Ideraabdullah, F. Y., Vigneau, S. and Bartolomei, M. S. (2008). Genomic imprinting mechanisms in mammals. *Mutat Res* **647**, 77-85.

Jacobson, R. H., Ladurner, A. G., King, D. S. and Tjian, R. (2000). Structure and function of a human TAFII250 double bromodomain module. *Science* **288**, 1422-5.
Jenuwein, T. and Allis, C. D. (2001). Translating the histone code. *Science* **293**, 1074-80.
Jessop, C. E., Watkins, R. H., Simmons, J. J., Tasab, M. and Bulleid, N. J. (2009). Protein disulphide isomerase family members show distinct substrate specificity: P5 is targeted to BiP client proteins. *J Cell Sci* **122**, 4287-95.
Jiang, J., Chan, Y. S., Loh, Y. H., Cai, J., Tong, G. Q., Lim, C. A., Robson, P., Zhong, S. and Ng, H. H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat Cell Biol* **10**, 353-60.
Jin, J., Cai, Y., Li, B., Conaway, R. C., Workman, J. L., Conaway, J. W. and Kusch, T. (2005). In and out: histone variant exchange in chromatin. *Trends Biochem Sci* **30**, 680-7.
Jorgensen, P. M., Graslund, S., Betz, R., Stahl, S., Larsson, C. and Hoog, C. (2001). Characterisation of the human APC1, the largest subunit of the anaphase-promoting complex. *Gene* **262**, 51-9.
Joseph, J., Mudduluru, G., Antony, S., Vashistha, S., Ajitkumar, P. and Somasundaram, K. (2004). Expression profiling of sodium butyrate (NaB)-treated cells: identification of regulation of genes related to cytokine signaling and cancer metastasis by NaB. *Oncogene* **23**, 6304-15.
Kamadurai, H. B., Souphron, J., Scott, D. C., Duda, D. M., Miller, D. J., Stringer, D., Piper, R. C. and Schulman, B. A. (2009). Insights into ubiquitin transfer cascades from a structure of a UbcH5B approximately ubiquitin-HECT(NEDD4L) complex. *Mol Cell* **36**, 1095-102.
Kargul, G. J., Dudekula, D. B., Qian, Y., Lim, M. K., Jaradat, S. A., Tanaka, T. S., Carter, M. G. and Ko, M. S. (2001). Verification and initial annotation of the NIA mouse 15K cDNA clone set. *Nat Genet* **28**, 17-8.
Karytinos, A., Forneris, F., Profumo, A., Ciossani, G., Battaglioli, E., Binda, C. and Mattevi, A. (2009). A novel mammalian flavin-dependent histone demethylase. *J Biol Chem* **284**, 17775-82.
Kaufman, P. D. and Rando, O. J. (2010). Chromatin as a potential carrier of heritable information. *Curr Opin Cell Biol* **22**, 284-90.
Kerppola, T. K. (2009). Polycomb group complexes--many combinations, many functions. *Trends Cell Biol* **19**, 692-704.
Kessel, M. and Gruss, P. (1991). Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* **67**, 89-104.
Kim, H., Kang, K. and Kim, J. (2009). AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2. *Nucleic Acids Res* **37**, 2940-50.
Kim, S. C., Sprung, R., Chen, Y., Xu, Y., Ball, H., Pei, J., Cheng, T., Kho, Y., Xiao, H., Xiao, L. et al. (2006). Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol Cell* **23**, 607-18.
Kimata, Y., Matsuyama, A., Nagao, K., Furuya, K., Obuse, C., Yoshida, M. and Yanagida, M. (2008). Diminishing HDACs by drugs or mutations promotes normal or abnormal sister chromatid separation by affecting APC/C and adherin. *J Cell Sci* **121**, 1107-18.
Klose, R. J., Kallin, E. M. and Zhang, Y. (2006). JmjC-domain-containing proteins and histone demethylation. *Nat Rev Genet* **7**, 715-27.

Klymenko, T., Papp, B., Fischle, W., Kocher, T., Schelder, M., Fritsch, C., Wild, B., Wilm, M. and Muller, J. (2006). A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. *Genes Dev* **20**, 1110-22.
Kmita, M. and Duboule, D. (2003). Organizing axes in time and space; 25 years of colinear tinkering. *Science* **301**, 331-3.
Kmita, M., van Der Hoeven, F., Zakany, J., Krumlauf, R. and Duboule, D. (2000). Mechanisms of Hox gene colinearity: transposition of the anterior Hoxb1 gene into the posterior HoxD complex. *Genes Dev* **14**, 198-211.
Knight, A. S., Notaridou, M. and Watson, R. J. (2009). A Lin-9 complex is recruited by B-Myb to activate transcription of G2/M genes in undifferentiated embryonal carcinoma cells. *Oncogene* **28**, 1737-47.
Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X. S. and Ahringer, J. (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat Genet* **41**, 376-81.
Kondo, T. and Duboule, D. (1999). Breaking colinearity in the mouse HoxD complex. *Cell* **97**, 407-17.
Kostrouchova, M., Kostrouch, Z. and Kostrouchova, M. (2007). Valproic acid, a molecular lead to multiple regulatory pathways. *Folia Biol (Praha)* **53**, 37-49.
Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* **128**, 693-705.
Kramer, O. H., Zhu, P., Ostendorff, H. P., Golebiewski, M., Tiefenbach, J., Peters, M. A., Brill, B., Groner, B., Bach, I., Heinzel, T. et al. (2003). The histone deacetylase inhibitor valproic acid selectively induces proteasomal degradation of HDAC2. *EMBO J* **22**, 3411-20.
Krumlauf, R. (1994). Hox genes in vertebrate development. *Cell* **78**, 191-201.
Ku, M., Koche, R. P., Rheinbay, E., Mendenhall, E. M., Endoh, M., Mikkelsen, T. S., Presser, A., Nusbaum, C., Xie, X., Chi, A. S. et al. (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet* **4**, e1000242.
Kuendgen, A. and Gattermann, N. (2007). Valproic acid for the treatment of myeloid malignancies. *Cancer* **110**, 943-54.
Kuhn, U. and Wahle, E. (2004). Structure and function of poly(A) binding proteins. *Biochim Biophys Acta* **1678**, 67-84.
Kundu, S., Horn, P. J. and Peterson, C. L. (2007). SWI/SNF is required for transcriptional memory at the yeast GAL gene cluster. *Genes Dev* **21**, 997-1004.
Kuzmichev, A., Jenuwein, T., Tempst, P. and Reinberg, D. (2004). Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. *Mol Cell* **14**, 183-93.
Lagarou, A., Mohd-Sarip, A., Moshkin, Y. M., Chalkley, G. E., Bezstarosti, K., Demmers, J. A. and Verrijzer, C. P. (2008). dKDM2 couples histone H2A ubiquitylation to histone H3 demethylation during Polycomb group silencing. *Genes Dev* **22**, 2799-810.
Laine, J. P., Singh, B. N., Krishnamurthy, S. and Hampsey, M. (2009). A physiological role for gene loops in yeast. *Genes Dev* **23**, 2604-9.
Lake, J., Rathjen, J., Remiszewski, J. and Rathjen, P. D. (2000). Reversible programming of pluripotent cell differentiation. *J Cell Sci* **113 (Pt 3)**, 555-66.
Landeira, D., Sauer, S., Poot, R., Dvorkina, M., Mazzarella, L., Jorgensen, H. F., Pereira, C. F., Leleu, M., Piccolo, F. M., Spivakov, M. et al. (2010). Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and

recruitment of PRC1 and RNA Polymerase II to developmental regulators. *Nat Cell Biol* **12**, 618-24.
Langston, A. W., Thompson, J. R. and Gudas, L. J. (1997). Retinoic acid-responsive enhancers located 3' of the Hox A and Hox B homeobox gene clusters. Functional analysis. *J Biol Chem* **272**, 2167-75.
Lanzuolo, C., Roure, V., Dekker, J., Bantignies, F. and Orlando, V. (2007). Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat Cell Biol* **9**, 1167-74.
Lee, B. M. and Mahadevan, L. C. (2009). Stability of histone modifications across mammalian genomes: implications for 'epigenetic' marking. *J Cell Biochem* **108**, 22-34.
Lee, D. Y., Hayes, J. J., Pruss, D. and Wolffe, A. P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* **72**, 73-84.
Lee, E. R., Murdoch, F. E. and Fritsch, M. K. (2007a). High histone acetylation and decreased polycomb repressive complex 2 member levels regulate gene specific transcriptional changes during early embryonic stem cell differentiation induced by retinoic acid. *Stem Cells* **25**, 2191-9.
Lee, K. K. and Workman, J. L. (2007). Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol* **8**, 284-95.
Lee, M. G., Villa, R., Trojer, P., Norman, J., Yan, K. P., Reinberg, D., Di Croce, L. and Shiekhattar, R. (2007b). Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* **318**, 447-50.
Lee, W., Tillo, D., Bray, N., Morse, R. H., Davis, R. W., Hughes, T. R. and Nislow, C. (2007c). A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet* **39**, 1235-44.
Leeb, M., Pasini, D., Novatchkova, M., Jaritz, M., Helin, K. and Wutz, A. (2010). Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev* **24**, 265-76.
Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T. and Peters, A. H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* **13**, 1192-200.
Lewis, E. B. (1978). A gene complex controlling segmentation in Drosophila. *Nature* **276**, 565-70.
Liang, G., Chan, M. F., Tomigahara, Y., Tsai, Y. C., Gonzales, F. A., Li, E., Laird, P. W. and Jones, P. A. (2002). Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Mol Cell Biol* **22**, 480-91.
Lim, D. A., Huang, Y. C., Swigut, T., Mirick, A. L., Garcia-Verdugo, J. M., Wysocka, J., Ernst, P. and Alvarez-Buylla, A. (2009). Chromatin remodelling factor Mll1 is essential for neurogenesis from postnatal neural stem cells. *Nature* **458**, 529-33.
Loenen, W. A. (2006). S-adenosylmethionine: jack of all trades and master of everything? *Biochem Soc Trans* **34**, 330-3.
Lopez-Serra, L. and Esteller, M. (2008). Proteins that bind methylated DNA and human cancer: reading the wrong words. *Br J Cancer* **98**, 1881-5.
Loyola, A., Bonaldi, T., Roche, D., Imhof, A. and Almouzni, G. (2006). PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Mol Cell* **24**, 309-16.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-60.
Maresca, T. J. and Heald, R. (2006). The long and the short of it: linker histone H1 is required for metaphase chromosome compaction. *Cell Cycle* **5**, 589-91.
Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C. L., Dynlacht, B. D. and Reinberg, D. (2008). Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol Cell* **32**, 503-18.
Margueron, R. and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343-9.
Marmorstein, R. (2001). Structure and function of histone acetyltransferases. *Cell Mol Life Sci* **58**, 693-703.
Marteijn, J. A., van der Meer, L. T., Smit, J. J., Noordermeer, S. M., Wissink, W., Jansen, P., Swarts, H. G., Hibbert, R. G., de Witte, T., Sixma, T. K. et al. (2009). The ubiquitin ligase Triad1 inhibits myelopoiesis through UbcH7 and Ubc13 interacting domains. *Leukemia* **23**, 1480-9.
Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* **78**, 7634-8.
Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A. A. et al. (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* **9**, 625-35.
Mateescu, B., England, P., Halgand, F., Yaniv, M. and Muchardt, C. (2004). Tethering of HP1 proteins to chromatin is relieved by phosphoacetylation of histone H3. *EMBO Rep* **5**, 490-6.
Matsunami, M., Sumiyama, K. and Saitou, N. (2010). Evolution of conserved non-coding sequences within the vertebrate Hox clusters through the two-round whole genome duplications revealed by phylogenetic footprinting analysis. *J Mol Evol* **71**, 427-36.
Meissner, A. (2010). Epigenetic modifications in pluripotent and differentiated cells. *Nat Biotechnol* **28**, 1079-88.
Mello, J. A., Sillje, H. H., Roche, D. M., Kirschner, D. B., Nigg, E. A. and Almouzni, G. (2002). Human Asf1 and CAF-1 interact and synergize in a repair-coupled nucleosome assembly pathway. *EMBO Rep* **3**, 329-34.
Mersfelder, E. L. and Parthun, M. R. (2006). The tale beyond the tail: histone core domain modifications and the regulation of chromatin structure. *Nucleic Acids Res* **34**, 2653-62.
Mesdjian, E., Ciesielski, L., Valli, M., Bruguerolle, B., Jadot, G., Bouyard, P. and Mandel, P. (1982). Sodium valproate: kinetic profile and effects on GABA levels in various brain areas of the rat. *Prog Neuropsychopharmacol Biol Psychiatry* **6**, 223-33.
Meshorer, E., Yellajoshula, D., George, E., Scambler, P. J., Brown, D. T. and Misteli, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* **10**, 105-16.
Metzger, E., Wissmann, M., Yin, N., Muller, J. M., Schneider, R., Peters, A. H., Gunther, T., Buettner, R. and Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* **437**, 436-9.

Meunier, H., Carraz, G., Neunier, Y., Eymard, P. and Aimard, M. (1963). [Pharmacodynamic properties of N-dipropylacetic acid.]. *Therapie* **18**, 435-8.
Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T. K., Koche, R. P. et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553-60.
Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M. and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631-42.
Mohn, F., Weber, M., Rebhan, M., Roloff, T. C., Richter, J., Stadler, M. B., Bibel, M. and Schubeler, D. (2008). Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* **30**, 755-66.
Moldovan, G. L., Pfander, B. and Jentsch, S. (2007). PCNA, the maestro of the replication fork. *Cell* **129**, 665-79.
Monarez, R. R., MacDonald, C. C. and Dass, B. (2007). Polyadenylation proteins CstF-64 and tauCstF-64 exhibit differential binding affinities for RNA polymers. *Biochem J* **401**, 651-8.
Morey, C., Da Silva, N. R., Kmita, M., Duboule, D. and Bickmore, W. A. (2008). Ectopic nuclear reorganisation driven by a Hoxb1 transgene transposed into Hoxd. *J Cell Sci* **121**, 571-7.
Mund, C. and Lyko, F. (2010). Epigenetic cancer therapy: Proof of concept and remaining challenges. *Bioessays* **32**, 949-57.
Mund, T. and Pelham, H. R. (2010). Regulation of PTEN/Akt and MAP kinase signaling pathways by the ubiquitin ligase activators Ndfip1 and Ndfip2. *Proc Natl Acad Sci U S A* **107**, 11429-34.
Murry, C. E. and Keller, G. (2008). Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132**, 661-80.
Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. and Roder, J. C. (1993). Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc Natl Acad Sci U S A* **90**, 8424-8.
Nau, H., Hauck, R. S. and Ehlers, K. (1991). Valproic acid-induced neural tube defects in mouse and human: aspects of chirality, alternative drug development, pharmacokinetics and possible mechanisms. *Pharmacol Toxicol* **69**, 310-21.
Negishi, M., Saraya, A., Miyagi, S., Nagao, K., Inagaki, Y., Nishikawa, M., Tajima, S., Koseki, H., Tsuda, H., Takasaki, Y. et al. (2007). Bmi1 cooperates with Dnmt1-associated protein 1 in gene silencing. *Biochem Biophys Res Commun* **353**, 992-8.
Nekrasov, M., Klymenko, T., Fraterman, S., Papp, B., Oktaba, K., Kocher, T., Cohen, A., Stunnenberg, H. G., Wilm, M. and Muller, J. (2007). Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. *EMBO J* **26**, 4078-88.
Nightingale, K. P., Gendreizig, S., White, D. A., Bradbury, C., Hollfelder, F. and Turner, B. M. (2007). Cross-talk between histone modifications in response to histone deacetylase inhibitors: MLL4 links histone H3 acetylation and histone H3K4 methylation. *J Biol Chem* **282**, 4408-16.
Nightingale, K. P., O'Neill, L. P. and Turner, B. M. (2006). Histone modifications: signalling receptors and potential elements of a heritable epigenetic code. *Curr Opin Genet Dev* **16**, 125-36.

Niwa, H., Miyazaki, J. and Smith, A. G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24**, 372-6.
Olins, A. L. and Olins, D. E. (1974). Spheroid chromatin units (v bodies). *Science* **183**, 330-2.
Pan, G., Li, J., Zhou, Y., Zheng, H. and Pei, D. (2006). A negative feedback loop of transcription factors that controls stem cell pluripotency and self-renewal. *FASEB J* **20**, 1730-2.
Pasini, D., Hansen, K. H., Christensen, J., Agger, K., Cloos, P. A. and Helin, K. (2008). Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. *Genes Dev* **22**, 1345-55.
Pasini, D., Malatesta, M., Jung, H. R., Walfridsson, J., Willer, A., Olsson, L., Skotte, J., Wutz, A., Porse, B., Jensen, O. N. et al. (2010). Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Res* **38**, 4958-69.
Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E. et al. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**, 517-27.
Pratt, W. B., Morishima, Y., Peng, H. M. and Osawa, Y. (2010). Proposal for a role of the Hsp90/Hsp70-based chaperone machinery in making triage decisions when proteins undergo oxidative and toxic damage. *Exp Biol Med (Maywood)* **235**, 278-89.
Pray-Grant, M. G., Daniel, J. A., Schieltz, D., Yates, J. R., 3rd and Grant, P. A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* **433**, 434-8.
Probst, A. V., Dunleavy, E. and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. *Nat Rev Mol Cell Biol* **10**, 192-206.
Ptashne, M. (2007). On the use of the word 'epigenetic'. *Curr Biol* **17**, R233-6.
Quivy, J. P., Gerard, A., Cook, A. J., Roche, D. and Almouzni, G. (2008). The HP1-p150/CAF-1 interaction is required for pericentric heterochromatin replication and S-phase progression in mouse cells. *Nat Struct Mol Biol* **15**, 972-9.
Rada-Iglesias, A., Enroth, S., Ameur, A., Koch, C. M., Clelland, G. K., Respuela-Alonso, P., Wilcox, S., Dovey, O. M., Ellis, P. D., Langford, C. F. et al. (2007). Butyrate mediates decrease of histone acetylation centered on transcription start sites and down-regulation of associated genes. *Genome Res* **17**, 708-19.
Ramsahoye, B. H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A. P. and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A* **97**, 5237-42.
Raney, B. J., Cline, M. S., Rosenbloom, K. R., Dreszer, T. R., Learned, K., Barber, G. P., Meyer, L. R., Sloan, C. A., Malladi, V. S., Roskin, K. M. et al. (2011). ENCODE whole-genome data in the UCSC genome browser (2011 update). *Nucleic Acids Res* **39**, D871-5.
Ransom, M., Dennehey, B. K. and Tyler, J. K. (2010). Chaperoning histones during DNA replication and repair. *Cell* **140**, 183-95.
Rastegar, M., Kobrossy, L., Kovacs, E. N., Rambaldi, I. and Featherstone, M. (2004). Sequential histone modifications at Hoxd4 regulatory regions distinguish anterior from posterior embryonic compartments. *Mol Cell Biol* **24**, 8090-103.
Raz, R., Lee, C. K., Cannizzaro, L. A., d'Eustachio, P. and Levy, D. E. (1999). Essential role of STAT3 for embryonic stem cell pluripotency. *Proc Natl Acad Sci U S A* **96**, 2846-51.

Reese, B. E., Bachman, K. E., Baylin, S. B. and Rountree, M. R. (2003). The methyl-CpG binding protein MBD1 interacts with the p150 subunit of chromatin assembly factor 1. *Mol Cell Biol* **23**, 3226-36.
Rice, J. C. and Allis, C. D. (2001). Histone methylation versus histone acetylation: new insights into epigenetic regulation. *Curr Opin Cell Biol* **13**, 263-73.
Ringrose, L. and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet* **38**, 413-43.
Ringrose, L. and Paro, R. (2007). Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development* **134**, 223-32.
Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., Goodnough, L. H., Helms, J. A., Farnham, P. J., Segal, E. et al. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**, 1311-23.
Rizzi, F. and Bettuzzi, S. (2010). The clusterin paradigm in prostate and breast carcinogenesis. *Endocr Relat Cancer* **17**, R1-17.
Robinson, P. J., An, W., Routh, A., Martino, F., Chapman, L., Roeder, R. G. and Rhodes, D. (2008). 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. *J Mol Biol* **381**, 816-25.
Rochette-Egly, C. and Germain, P. (2009). Dynamic and combinatorial control of gene expression by nuclear retinoic acid receptors (RARs). *Nucl Recept Signal* **7**, e005.
Rodnina, M. V. and Wintermeyer, W. (2009). Recent mechanistic insights into eukaryotic ribosomes. *Curr Opin Cell Biol* **21**, 435-43.
Roh, T. Y., Cuddapah, S., Cui, K. and Zhao, K. (2006). The genomic landscape of histone modifications in human T cells. *Proc Natl Acad Sci U S A* **103**, 15782-7.
Sacher, M., Di Bacco, A., Lunin, V. V., Ye, Z., Wagner, J., Gill, G. and Cygler, M. (2005). The crystal structure of CREG, a secreted glycoprotein involved in cellular growth and differentiation. *Proc Natl Acad Sci U S A* **102**, 18326-31.
Sanchez, R. and Zhou, M. M. (2009). The role of human bromodomains in chromatin biology and gene transcription. *Curr Opin Drug Discov Devel* **12**, 659-65.
Sarma, K., Margueron, R., Ivanov, A., Pirrotta, V. and Reinberg, D. (2008). Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo. *Mol Cell Biol* **28**, 2718-31.
Scheuermann, J. C., de Ayala Alonso, A. G., Oktaba, K., Ly-Hartig, N., McGinty, R. K., Fraterman, S., Wilm, M., Muir, T. W. and Muller, J. (2010). Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. *Nature* **465**, 243-7.
Schuettengruber, B. and Cavalli, G. (2009). Recruitment of polycomb group complexes and their role in the dynamic regulation of cell fate choice. *Development* **136**, 3531-42.
Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B. and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. *Cell* **128**, 735-45.
Selvi, B. R., Cassel, J. C., Kundu, T. K. and Boutillier, A. L. (2010). Tuning acetylation levels with HAT activators: therapeutic strategy in neurodegenerative diseases. *Biochim Biophys Acta* **1799**, 840-53.
Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J. R., Wu, C. T., Bender, W. and Kingston, R. E. (1999). Stabilization of chromatin structure by PRC1, a Polycomb complex. *Cell* **98**, 37-46.
Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T. A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K. et al. (2007). The SRA protein Np95

mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**, 908-12.
Shembade, N., Parvatiyar, K., Harhaj, N. S. and Harhaj, E. W. (2009). The ubiquitin-editing enzyme A20 requires RNF11 to downregulate NF-kappaB signalling. *EMBO J* **28**, 513-22.
Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A. and Casero, R. A. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941-53.
Shiba, K., Suzuki, N., Shigesada, K., Namba, Y., Schimmel, P. and Noda, T. (1994). Human cytoplasmic isoleucyl-tRNA synthetase: selective divergence of the anticodon-binding domain and acquisition of a new structural unit. *Proc Natl Acad Sci U S A* **91**, 7435-9.
Shibahara, K. and Stillman, B. (1999). Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. *Cell* **96**, 575-85.
Shogren-Knaak, M., Ishii, H., Sun, J. M., Pazin, M. J., Davie, J. R. and Peterson, C. L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**, 844-7.
Silva, J. and Smith, A. (2008). Capturing pluripotency. *Cell* **132**, 532-6.
Simon, J. A. and Kingston, R. E. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol* **10**, 697-708.
Simon, J. A. and Tamkun, J. W. (2002). Programming off and on states in chromatin: mechanisms of Polycomb and trithorax group complexes. *Curr Opin Genet Dev* **12**, 210-8.
Sing, A., Pannell, D., Karaiskakis, A., Sturgeon, K., Djabali, M., Ellis, J., Lipshitz, H. D. and Cordes, S. P. (2009). A vertebrate Polycomb response element governs segmentation of the posterior hindbrain. *Cell* **138**, 885-97.
Solomon, E., Li, H., Muggy, S. D., Syta, E. and Zolkiewska, A. (2010). The role of SnoN in transforming growth factor beta1-induced expression of metalloprotease-disintegrin ADAM12. *J Biol Chem* **285**, 21969-77.
Solter, D. (2006). From teratocarcinomas to embryonic stem cells and beyond: a history of embryonic stem cell research. *Nat Rev Genet* **7**, 319-27.
Solter, D. and Knowles, B. B. (1978). Monoclonal antibody defining a stage-specific mouse embryonic antigen (SSEA-1). *Proc Natl Acad Sci U S A* **75**, 5565-9.
Soshnikova, N. and Duboule, D. (2009). Epigenetic temporal control of mouse Hox genes in vivo. *Science* **324**, 1320-3.
Spitz, F., Gonzalez, F. and Duboule, D. (2003). A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell* **113**, 405-17.
Spivakov, M. and Fisher, A. G. (2007). Epigenetic signatures of stem-cell identity. *Nat Rev Genet* **8**, 263-71.
Spotswood, H. T. and Turner, B. M. (2002). An increasingly complex code. *J Clin Invest* **110**, 577-82.
Spriet, L. L. and Heigenhauser, G. J. (2002). Regulation of pyruvate dehydrogenase (PDH) activity in human skeletal muscle during exercise. *Exerc Sport Sci Rev* **30**, 91-5.
Stead, E., White, J., Faast, R., Conn, S., Goldstone, S., Rathjen, J., Dhingra, U., Rathjen, P., Walker, D. and Dalton, S. (2002). Pluripotent cell division cycles are driven by ectopic Cdk2, cyclin A/E and E2F activities. *Oncogene* **21**, 8320-33.
Stewart, M. (2007). Ratcheting mRNA out of the nucleus. *Mol Cell* **25**, 327-30.

Stewart, M. D., Li, J. and Wong, J. (2005). Relationship between histone H3 lysine 9 methylation, transcription repression, and heterochromatin protein 1 recruitment. *Mol Cell Biol* **25**, 2525-38.
Stock, J. K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A. G. and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nat Cell Biol* **9**, 1428-35.
Szutorisz, H., Canzonetta, C., Georgiou, A., Chow, C. M., Tora, L. and Dillon, N. (2005). Formation of an active tissue-specific chromatin domain initiated by epigenetic marking at the embryonic stem cell stage. *Mol Cell Biol* **25**, 1804-20.
Taft, R. J., Pheasant, M. and Mattick, J. S. (2007). The relationship between non-protein-coding DNA and eukaryotic complexity. *Bioessays* **29**, 288-99.
Tagami, H., Ray-Gallet, D., Almouzni, G. and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116**, 51-61.
Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-76.
Takai, N. and Narahara, H. (2008). Array-based approaches for the identification of epigenetic silenced tumor suppressor genes. *Curr Genomics* **9**, 22-4.
Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. and Patel, D. J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* **14**, 1025-40.
Terrenoire, E., McRonald, F., Halsall, J. A., Page, P., Illingworth, R. S., Taylor, A. M., Davison, V., O'Neill, L. P. and Turner, B. M. (2010). Immunostaining of modified histones defines high-level features of the human metaphase epigenome. *Genome Biol* **11**, R110.
Thoma, F., Koller, T. and Klug, A. (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J Cell Biol* **83**, 403-27.
Thorne, A. W., Kmiciek, D., Mitchelson, K., Sautiere, P. and Crane-Robinson, C. (1990). Patterns of histone acetylation. *Eur J Biochem* **193**, 701-13.
Tian, L. M. and Alkadhi, K. A. (1994). Valproic acid inhibits the depolarizing rectification in neurons of rat amygdala. *Neuropharmacology* **33**, 1131-8.
Tie, F., Banerjee, R., Stratton, C. A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M. O., Scacheri, P. C. and Harte, P. J. (2009). CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. *Development* **136**, 3131-41.
Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**, 4350-4.
Tsai, M. C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., Shi, Y., Segal, E. and Chang, H. Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**, 689-93.
Tse, C., Sera, T., Wolffe, A. P. and Hansen, J. C. (1998). Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. *Mol Cell Biol* **18**, 4629-38.
Tumpel, S., Cambronero, F., Ferretti, E., Blasi, F., Wiedemann, L. M. and Krumlauf, R. (2007). Expression of Hoxa2 in rhombomere 4 is regulated by a conserved cross-regulatory mechanism dependent upon Hoxb1. *Dev Biol* **302**, 646-60.

Turner, B. M. (2009). Epigenetic responses to environmental change and their evolutionary implications. *Philos Trans R Soc Lond B Biol Sci* **364**, 3403-18.
Turner, B. M. and Fellows, G. (1989). Specific antibodies reveal ordered and cell-cycle-related use of histone-H4 acetylation sites in mammalian cells. *Eur J Biochem* **179**, 131-9.
Turner, B. M., O'Neill, L. P. and Allan, I. M. (1989). Histone H4 acetylation in human cells. Frequency of acetylation at different sites defined by immunolabeling with site-specific antibodies. *FEBS Lett* **253**, 141-5.
Tyagi, S., Chabes, A. L., Wysocka, J. and Herr, W. (2007). E2F activation of S phase promoters via association with HCF-1 and the MLL family of histone H3K4 methyltransferases. *Mol Cell* **27**, 107-19.
van der Hoeven, F., Zakany, J. and Duboule, D. (1996). Gene transpositions in the HoxD complex reveal a hierarchy of regulatory controls. *Cell* **85**, 1025-35.
VerMilyea, M. D., O'Neill, L. P. and Turner, B. M. (2009). Transcription-independent heritability of induced histone modifications in the mouse preimplantation embryo. *PLoS One* **4**, e6086.
Vettese-Dadey, M., Grant, P. A., Hebbes, T. R., Crane- Robinson, C., Allis, C. D. and Workman, J. L. (1996). Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *EMBO J* **15**, 2508-18.
Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J. M. et al. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**, 871-4.
Voncken, J. W., Schweizer, D., Aagaard, L., Sattler, L., Jantsch, M. F. and van Lohuizen, M. (1999). Chromatin-association of the Polycomb group protein BMI1 is cell cycle-regulated and correlates with its phosphorylation status. *J Cell Sci* **112 (Pt 24)**, 4627-39.
Wang, Z., Zang, C., Cui, K., Schones, D. E., Barski, A., Peng, W. and Zhao, K. (2009). Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**, 1019-31.
Watanabe, Y. and Maekawa, M. (2010). Methylation of DNA in cancer. *Adv Clin Chem* **52**, 145-67.
Widom, J. (1986). Physicochemical studies of the folding of the 100 A nucleosome filament into the 300 A filament. Cation dependence. *J Mol Biol* **190**, 411-24.
Woo, C. J., Kharchenko, P. V., Daheron, L., Park, P. J. and Kingston, R. E. (2010). A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell* **140**, 99-110.
Wysocka, J., Myers, M. P., Laherty, C. D., Eisenman, R. N. and Herr, W. (2003). Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. *Genes Dev* **17**, 896-911.
Yamauchi, T., Yamauchi, J., Kuwata, T., Tamura, T., Yamashita, T., Bae, N., Westphal, H., Ozato, K. and Nakatani, Y. (2000). Distinct but overlapping roles of histone acetylase PCAF and of the closely related PCAF-B/GCN5 in mouse embryogenesis. *Proc Natl Acad Sci U S A* **97**, 11303-6.
Yang, B. and Kumar, S. (2010). Nedd4 and Nedd4-2: closely related ubiquitin-protein ligases with distinct physiological functions. *Cell Death Differ* **17**, 68-77.

Yang, X. J. and Seto, E. (2007). HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* **26**, 5310-8.
Yao, T. P., Oh, S. P., Fuchs, M., Zhou, N. D., Ch'ng, L. E., Newsome, D., Bronson, R. T., Li, E., Livingston, D. M. and Eckner, R. (1998). Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* **93**, 361-72.
Yap, K. L., Li, S., Munoz-Cabello, A. M., Raguz, S., Zeng, L., Mujtaba, S., Gil, J., Walsh, M. J. and Zhou, M. M. (2010). Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* **38**, 662-74.
Youn, H. D., Grozinger, C. M. and Liu, J. O. (2000). Calcium regulates transcriptional repression of myocyte enhancer factor 2 by histone deacetylase 4. *J Biol Chem* **275**, 22563-7.
Zacharioudakis, I., Gligoris, T. and Tzamarias, D. (2007). A yeast catabolic enzyme controls transcriptional memory. *Curr Biol* **17**, 2041-6.
Zee, B. M., Levin, R. S., Dimaggio, P. A. and Garcia, B. A. (2010). Global turnover of histone post-translational modifications and variants in human cells. *Epigenetics Chromatin* **3**, 22.
Zhang, J., Xu, F., Hashimshony, T., Keshet, I. and Cedar, H. (2002). Establishment of transcriptional competence in early and late S phase. *Nature* **420**, 198-202.
Zhang, Z., Shibahara, K. and Stillman, B. (2000). PCNA connects DNA replication to epigenetic inheritance in yeast. *Nature* **408**, 221-5.
Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J. and Lee, J. T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322**, 750-6.
Zhou, H., Hu, H. and Lai, M. (2010). Non-coding RNAs and their epigenetic regulatory mechanisms. *Biol Cell* **102**, 645-55.
Zilberman, D., Coleman-Derr, D., Ballinger, T. and Henikoff, S. (2008). Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* **456**, 125-9.