PROSPECT
Journal of Cellular Biochemistry 108:22–34 (2009)

Journal of Cellular Biochemistry

Stability of Histone Modifications Across Mammalian Genomes: Implications for 'Epigenetic' Marking

Benjamin M. Lee and Louis C. Mahadevan*

Nuclear Signalling Laboratory, Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK

ABSTRACT

The combination of chromatin immunoprecipitation (ChIP) with microarray analysis (ChIP-chip) or high-throughput sequencing (ChIP-seq and ChIP-SAGE) has provided maps of a wide variety of site-specific histone modifications across mammalian genomes in various cell types. Although distinct genomic regions and functional elements have been strongly associated with specific histone modifications, an overwhelming number of combinatorial patterns have also been observed across the genome. While peaks of enrichment in ChIP-chip and ChIP-seq data may suggest stable and predictive 'landmarks' across the genomic landscape, studies from transcribed genes indicate a more dynamic model of how these data may be interpreted. In light of such studies, which show highly dynamic methylation, acetylation and phosphorylation of histone H3 during gene transcription, we consider the extent to which genome-wide maps of chromatin state could be interpreted as 'snapshots' of heterogeneous profiles deriving from dynamic modification processes. Rather than acting as static 'epigenetic' landmarks, histone modifications may function as dynamic and transient operational marks supporting specific steps in diverse processes throughout the mammalian genome. J. Cell. Biochem. 108: 22–34, 2009.

KEY WORDS: HISTONE MODIFICATIONS IN MAMMALS; CHROMATIN IN MAMMALIAN GENOMES; QUANTITATION OF ChIP ASSAYS

DNA encoding the eukaryotic genome is packaged and regulated through intimate association with histones to form nucleosomes, the repeating unit of chromatin. The canonical nucleosome core particle consists of 146 bp of DNA wrapped around an octamer of the core histones H2A, H2B, H3 and H4 [Luger et al., 1997]. Within chromatin, nucleosomes exhibit considerable structural and compositional diversity, achieved through histone variant incorporation, alternative nucleosomal structures, DNA methylation and a vast array of post-translational histone modifications [Sarma and Reinberg, 2005; Klose and Bird, 2006; Kouzarides, 2007; Zlatanova et al., 2009]. Since Vincent Allfrey's discovery of histone methylation and acetylation over 40 years ago [Allfrey et al., 1964], such modifications (presently including phosphorylation, ubiquitinylation, sumoylation, biotinylation, ADP-ribosylation, prolyl isomerisation and tail clipping) have been strongly implicated in virtually all genomic processes [Kouzarides, 2007]. Various functions proposed for histone modification now fall into three classes: (i) to alter biophysical properties of nucleosomes; (ii) to act as a 'histone code', relying on combinatorial patterns and antagonistic or synergistic cross-talk that is decoded by specific effector proteins [Strahl and Allis, 2000; Taverna et al., 2007]; or (iii) to act akin to signal transduction pathways, utilising redundancy and reinforcement by feedback loops [Schreiber and Bernstein, 2002].

Exploiting the compact nature of the yeast genome, chromatin immunoprecipitation (ChIP)-coupled microarray analyses provided first indications of the distribution and implied function of histone modifications on global scale in a unicellular eukaryote [Bernstein et al., 2002; Pokholok et al., 2005]. In higher eukaryotes including mammals, the subject of this review, multicellularity and the evolution of complex embryological and differentiation patterns may require evolution of more sophisticated modes of nucleosomal and genomic regulation, including new modes of controlling deposition and function of histone modifications. Association of specific modifications with particular patterns of gene expression has promoted their potential involvement in differentiation and epigenetic inheritance, whereby stable histone modifications define heritable patterns of gene expression and thus cellular identity, but

*Correspondence to: Prof. Louis C. Mahadevan, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK.
E-mail: louis.mahadevan@bioch.ox.ac.uk
Received 22 May 2009; Accepted 26 May 2009 • DOI 10.1002/jcb.22250 • © 2009 Wiley-Liss, Inc.
Published online 21 July 2009 in Wiley InterScience (www.interscience.wiley.com).

the hypothesis that histone modifications function as heritable epigenetic marks remains contentious [Bird, 2007]. Mapping of the mammalian 'epigenome'—the spatial and temporal complement of genome-wide chromatin modifications in different cell types and environmental conditions—is an important step towards resolving this issue. Recent advances in massively parallel sequencing now facilitate comprehensive mapping of a wide-variety of modifications across larger mammalian genomes in different cell types (Table I) [Barski et al., 2007; Mikkelsen et al., 2007; Ku et al., 2008; Wang et al., 2008; Cui et al., 2009; Rosenfeld et al., 2009]. This review is focused on correlative and predictive descriptions of histone modifications associated with specific genomic elements based on genome-wide profiling in mammalian systems. Interpretation of these global maps is seriously complicated by ambiguity in quantitative information and potential heterogeneity of experimental systems under analysis. We discuss the possibility that such ChIP profiles may also be interpreted as indicative of micro-heterogeneity derived from dynamic modification processes. One consequence of this interpretation is that rather than acting as static epigenetic landmarks in the genomic landscape, histone modifications may act as dynamic and transient operational marks, exemplified by those described at poised and actively transcribed genes. Precise quantitative analysis of the absolute levels and dynamics of these modifications is required to judge the likelihood that histone modifications can act epigenetically to program different loci in different cell types.

TABLE I. Selected Examples of Systems Utilised to Map Histone Modifications Across Mammalian Genomes

Reference	Cell-type	Technique	Histone modifications		
Barski et al. [2007]	Human resting $CD4^+$ T-cells	Native, MNase-digested ChIP-seq	H2BK5me1 H3R2me1 H3R2me2a H3K4me1 H3K4me2 H3K4me3 H3K9me1	H3K9me2 H3K9me3 H3K27me1 H3K27me2 H3K27me3 H3K36me1 H3K36me3	H3K79me1 H3K79me2 H3K79me3 H4K20me1 H4K20me3
Bernstein et al. [2005]	Human lung/gum/toe/foreskin fibroblasts; hepatoma, melanoma and carcinoma cells; MLFs	X-ChIP-chip	H3K4me2 H3K4me3	H3K9ac(K14ac)*	
Bernstein et al. [2006]	mESCs, MEFs, MLFs, mouse myoblast and neuroblastoma cells	X-ChIP-chip Native, MNase-digested ChIP-chip	H3K4me3	H3K27me3	
Boyer et al. [2006]	mESCs	X-ChIP-chip	H3K27me3		
Cui et al. [2009]	Purified human $CD133^+$ HSPCs and $CD36^+$ erythrocyte precursors	Native, MNase-digested ChIP-seq	H3K4me1 H3K4me3 H3K9me1	H3K9me3 H3K27me1 H3K27me3	H3K36me3 H4K20me1
Guenther et al. [2007]	hESCs, human leukaemia cells, human hepatocytes	X-ChIP-chip	H3K4me3 H3K36me3 H3K79me2	H3K9ac(K14ac)*	
Heintzman et al. [2007]	HeLa	X-ChIP-chip	H3K4me1 H3K4me2 H3K4me3	H3K9ac(K14ac)* H4ac	
Heintzman et al. [2009]	hESCs, HeLa, human erythroleukaemia cells, immortalised lymphoblast, embryonic lung fibroblasts	X-ChIP-chip	H3K4me1 H3K4me2	H3K4me3 H3K9ac	H3K18ac H3K27ac
Koch et al. [2007]	HeLa, human immortalised lymphoblasts, fetal lung fibroblasts	X-ChIP-chip	H3K4me1 H3K4me2 H3K4me3	H3K9ac(K14ac)* H4ac	
Ku et al. [2008]	mESCs, hESCs	X-ChIP-seq	H3K4me3	H3K27me3	H3K36me3
Mikkelsen et al. [2007]	mESCs, MEFs, mouse NPCs	X-ChIP-seq	H3K4me3 H3K9me3 H3K27me3	H3K36me3 H4K20me3	
Orford et al. [2008]	mESCs, mouse HSPCs	X-ChIP-chip	H3K4me2	H3K4me3	H4ac
Pan et al. [2007]	hESCs	X-ChIP-chip	H3K4me3	H3K27me3	
Roh et al. [2005]	Human resting $CD4^+$ T-cells	X-ChIP-SAGE	H3K9ac(K14ac)*		
Roh et al. [2006]	Human resting $CD4^+$ T-cells	X-ChIP-SAGE	H3K4me3 H3K27me3	H3K9ac(K14ac)*	
Wang et al. [2008]	Human resting $CD4^+$ T-cells	Native, MNase-digested ChIP-seq	H2AK5ac H2AK9ac H2BK5ac H2BK12ac H2BK20ac H2BK120ac H3K4ac	H3K9ac H3K14ac H3K18ac H3K23ac H3K27ac H3K36ac H3K79me1	H3K79me2 H3K79me3 H4K5ac H4K8ac H4K12ac H4K16ac H4K91ac
Wei et al. [2009]	Human $CD4^+$ T-cells, Th1, Th2, T-regulatory	Native, MNase-digested ChIP-seq	H3K4me3	H3K27me3	
Wen et al. [2009]	Human placenta, mESCs	Native, MNase-digested ChIP-seq	H3K9me2	H3K9me3	H3K27me3
Xie et al. [2009]	hESCs, HeLa, human retinal pigment epithelial cells, embryonic kidney cells, human foreskin fibroblasts	X-ChIP-chip	H3K4me3 H3K56ac	H3K27me3	H3K9ac
Zhao et al. [2007]	hESCs	ChIP-PET	H3K4me3	H3K27me3	

ESC, embryonic stem cell; HSPC, haematopoietic stem/progenitor cell; MEF, mouse embryonic fibroblast; MLF, mouse lung fibroblast; MNase, micrococcal nuclease; NPC, neural progenitor cell; PET, paired-end ditag; SAGE, serial analysis of gene expression; X-ChIP, formaldehyde-crosslinked, sonicated chromatin immunoprecipitation; histone modifications denoted by the Brno nomenclature [Turner, 2005]; H4ac is a pan-acetyl antibody against H4K5acK8acK12acK16ac.
*Upstate 06-599 antibody raised against H3K9acK14ac shown to be predominantly specific to K9ac exclusively [Edmondson et al., 2002; Wang et al., 2008].

HISTONE MODIFICATION MAPPING ACROSS MAMMALIAN GENOMES: TECHNIQUES AND THE IMPORTANCE OF QUANTITATION

The quintessential technique to study association of specific histone modifications with DNA is chromatin immunoprecipitation (ChIP) [Hebbes et al., 1988]. This involves recovery and analysis of DNA from chromatin fragments immunoselected with antibodies specific to the modification of interest. Raising such antibodies requires knowledge of their location on histone tails; mass spectrometry is thus an important tool to identify modifications and their co-localisation with additional marks on the same histone tail, the presence of which may potentially affect antibody recognition by occlusion [Clayton et al., 2006]. Quantification of specific DNA sequences present in immunoprecipitated chromatin has traditionally relied on radioactive or real-time (RT) PCR using primers spanning defined amplicons [Thomson et al., 2001; Hazzalin and Mahadevan, 2005; Vakoc et al., 2006; Kim et al., 2007; Edmunds et al., 2008]. Information and interpretations from ChIP-based studies are ultimately dictated by several fundamental ChIP parameters: issues relating to cell population heterogeneity, preparation of chromatin fragments, antibody quality, data presentation and their interpretation have been discussed at length previously [Clayton et al., 2006]. A recent publication highlights this issue: varying experimental conditions markedly affects recovery of gene sequences associated with H3K79-methylated nucleosomes [Steger et al., 2008], with consequences for previous interpretations of their distribution in mammalian genomes [Barski et al., 2007]. Nevertheless, ChIP coupled to RT-PCR is powerful and provides some of the most quantitative and high-resolution maps available, usefully expressed as percentages of total input material indicating absolute proportions of modified nucleosomes at any position. However, this approach, requiring use of pairs of PCR primers for each position, is restricted to analysis of precise predefined regions within mammalian genomes.

Larger scale genome-wide and chromosomal studies have utilised three complementary approaches to map histone modifications (Table I): (i) linear-amplification of immunoprecipitated DNA and hybridisation to tiling oligonucleotide microarrays (ChIP-chip) [Bernstein et al., 2006; Heintzman et al., 2009]; (ii) direct sequencing of concatenated DNA sequences by ChIP-SAGE (serial analysis of gene expression) [Roh et al., 2005; Zhao et al., 2007]; or (iii) direct massively parallel sequencing of short fragments of ChIP DNA (ChIP-seq) [Barski et al., 2007; Mikkelsen et al., 2007]. Advantages and drawbacks of these techniques have been comprehensively evaluated elsewhere [Buck and Lieb, 2004; Barski and Zhao, 2009]. A critical issue in genomic and chromosomal maps generated by these methods currently relates to their quantification and whether absolute levels of modification or the proportion of modified nucleosomes at any position can be inferred from such data (discussed further below). Although fragment densities in ChIP-seq positively correlate with fold-enrichment determined by RT-PCR [Mikkelsen et al., 2007], presentation of data as peaks or islands of enrichment signals obscures the precise proportion of nucleosomes at a given genomic position that carries a specific modification.

In virtually all ChIP studies, the nature of chromatin at regions described with overlapping peaks of different histone modifications needs to be interpreted with caution. Overlapping peaks of enrichment do not necessarily indicate co-localisation of the modifications on the same histone tail or even the same nucleosome, and may equally be attributable to the different scenarios depicted in Figure 1d. Furthermore, unless an antibody is demonstrated to specifically immunoprecipitate virtually all of its cognate epitope from input chromatin, comparisons of relative levels of histone modifications at specific positions are invalid. With poor immunoprecipitation efficiency, it is impossible to assess the actual contribution of real quantitative changes in histone modification (over changes in epitope accessibility, for example) against changes in: (i) fold-enrichment; (ii) quantitative differences in histone modification between different loci under different conditions; or (iii) different modifications at the same genomic locations. Quantitative analysis of the precise fraction of nucleosomes bearing a particular modification at a given position can indicate the existence of microheterogeneity and dynamic turnover (discussed further in Immediate-Early Gene Induction Highlights the Dynamics and Microheterogeneity of Histone Modification Section) [Thomson et al., 2001; Edmunds et al., 2008]. These properties of chromatin are generally not available from current genome-wide data, which largely portray histone modifications as static landmarks within particular genomic and cellular contexts [Barski et al., 2007; Guenther et al., 2007; Koch et al., 2007; Mikkelsen et al., 2007; Guttman et al., 2009]. Although data from these studies produce important qualitative information on global distribution patterns, ambiguity in quantitation prevents deeper insights into their precise amounts and relative significance at any position. In the Histone Modifications as Dynamic Operational Marks in Genomic Processes Section, we argue that quantitative analysis of histone modification produces the clearest insights into interpretations of chromatin modifications as either stable epigenetic marks or short-term, dynamic post-translational modifications.

HISTONE MODIFICATIONS AS LANDMARKS OF MAMMALIAN GENOMIC AND CELLULAR IDENTITY

The various associations of histone modifications described in this section should be viewed with the cautions described above, especially with regard to quantitative issues. Using data from mapping studies, distinctive sets of modifications have been associated with different functional regions along a mammalian chromosome (Fig. 1). Two general classes of spatial distribution can be distinguished: (i) broad domains spanning several kb, containing large arrays of modified nucleosomes (Fig. 1a); and (ii) punctate peaks within 1 kb, corresponding to a few highly localised modified nucleosomes (Fig. 1b,c).

DISTINCT SETS OF MODIFICATIONS DEMARCATE NON-GENIC CHROMOSOMAL FEATURES: ANALYSIS OF TELOMERES, CENTROMERES AND REPETITIVE DNA

A large proportion of mammalian genomes is composed of non-genic, repetitive DNA sequence, derived from DNA transposons,

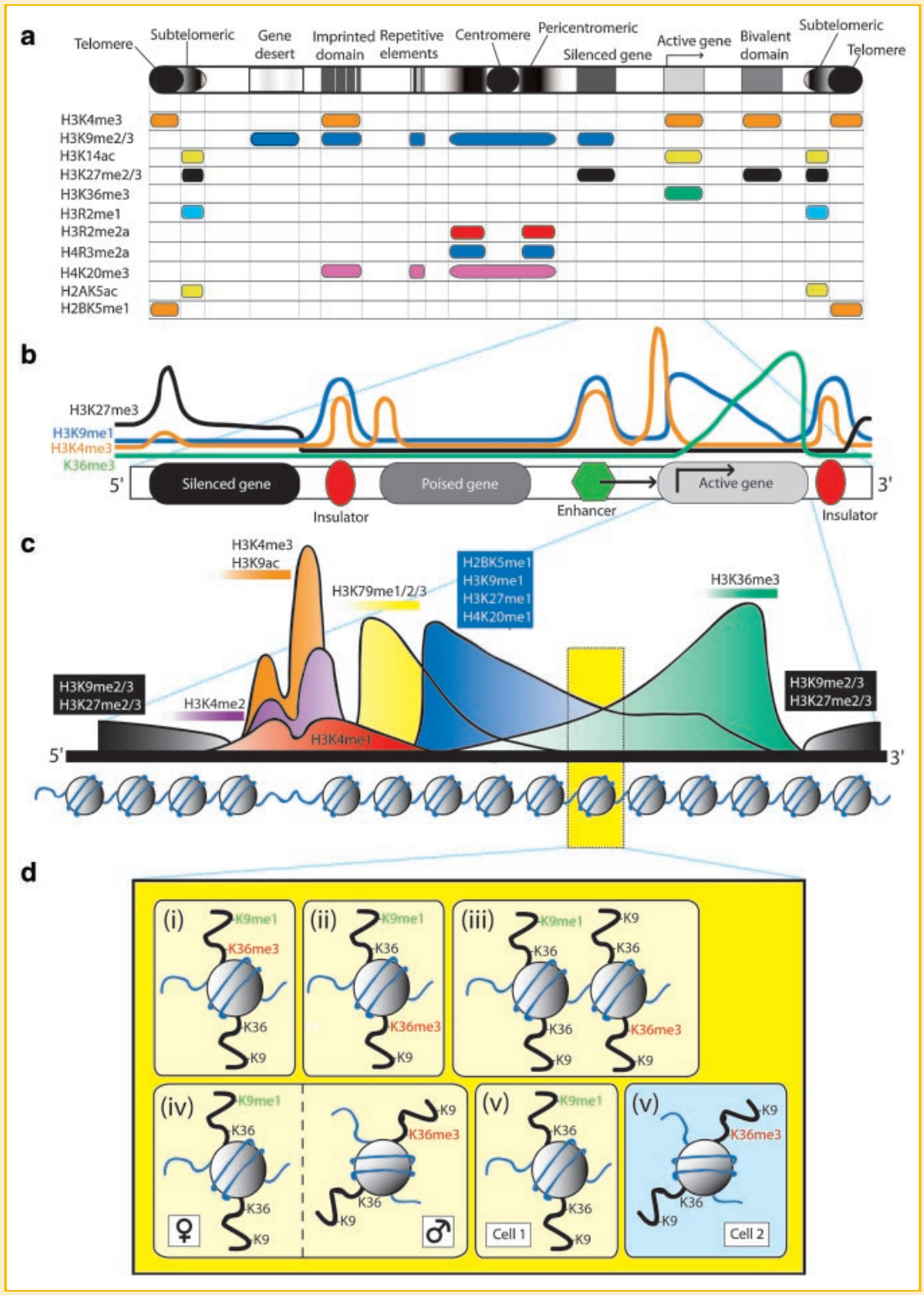

Fig. 1. Histone modifications across a typical mammalian chromosome. a: Schematic representation of histone modifications across a typical mammalian chromosome. Histone modifications are designated by the Brno nomenclature: acetylation (ac), mono/di/trimethylation (me1/me2/me3), asymmetrical dimethylation (me2a); for example H3K4me3 designates trimethylation of lysine 4 on histone H3. See main text for conflicting views on H3K9me2/3 at centromeric regions. b: Composite profile of typical histone modifications associated with active, poised, silenced genes and associated regulatory elements. See main text, however, for conflicting views on modifications at enhancers. c: Composite profile of typical histone modifications associated with an actively transcribed gene. Distributions derived from data presented in Bernstein et al. [2006], Barski et al. [2007], Heintzman et al. [2007], Mikkelsen et al. [2007], Wang et al. [2008], Heintzman et al. [2009], Rosenfeld et al. [2009]. d: Overlapping peaks of modification (H3K9me1 and H3K36me3, for example) may result from modification of: (i) the same histone; (ii) different histones in the same nucleosome; (iii) different nucleosomes on the same chromatin fragment; (iv) nucleosomes on different chromosomes (alleles) within the same cell; or (v) nucleosomes in different cells.

retrotransposons, satellite repeats and structures that form telomeres and centromeres [Lander et al., 2001]. These features are largely packaged into heterochromatin, traditionally associated with repressive modifications H3K9me2/3, H3K27me3 and H4K20me3 [Martens et al., 2005; Barski et al., 2007; Mikkelsen et al., 2007]. Although direct genome-wide mapping of repetitive elements is hindered by cross-hybridisation, alignment to consensus repeat sequences enables discrimination between centromeric and telomeric regions [Rosenfeld et al., 2009]. Such analysis reveals that distinct histone modification patterns are associated with different structures across a typical mammalian chromosome (Fig. 1a).

Centromeres are composed of higher-order repeats of a series of satellite monomers. ChIP-seq analysis suggests they are enriched with H3K9me2/3 and H4K20me3 (Fig. 1a) [Rosenfeld et al., 2009]. This conflicts, however, with an immunofluorescence study that showed H3K9me2/3 is excluded from centromeric subdomains containing the CENP-A histone H3 variant, which are associated with H3K4me2 (but not H3K4me3) and hypoacetylated H3/H4 [Sullivan and Karpen, 2004]. Pericentromeric chromatin is also gene-poor, rich in repetitive satellite elements and interspersed with long and short interspersed nuclear elements (LINEs and SINEs, respectively). It contains H3K9me2/3 and H4K20me3 but is also significantly associated with H3R2me2a and H4R3me2a (Fig. 1a) [Rosenfeld et al., 2009]. H3K4me1/2/3 is excluded from pericentromeric regions, consistent with the presence of H3R2me2a which has been reported to inhibit H3K4 methylation and binding of H3K4me3-effectors [Guccione et al., 2007; Iberg et al., 2008].

At telomeres, ChIP-seq studies in mouse embryonic stem cells (mESCs) and differentiated cells showed greatest enrichments of H3K9me3 and H4K20me3 [Mikkelsen et al., 2007]. Intriguingly, this study showed that these regions are five-fold more enriched in K9me3 in mESCs than in differentiated cells [Mikkelsen et al., 2007]. However, analysis in human $CD4^+$ T-cells by alignment to telomeric or centromeric consensus sequences suggests H3K9me3 and H4K20me3 are predominantly contributed from centromeric chromatin [Rosenfeld et al., 2009]. Telomeric and subtelomeric regions lack both these modifications and are distinguishible: telomeres are most significantly associated with H2BK5me1 and H3K4me3, while subtelomeric regions are marked with H3K27me2/3, H2AK5ac, H3K14ac and H3R2me1 (Fig. 1a) [Rosenfeld et al., 2009]. The lack of H3K9me3 and H4K20me3, typically associated with heterochromatin, and presence of marks associated with active genes (H3K4me3, H3K14ac, H2BK5me1, H2AK5ac) suggests that a different mode of repression operates at telomeres compared to centromeres.

Other repetitive elements across the mammalian genome are repressed through heterochromatic mechanisms. Significant enrichments of H3K9me3 and H4K20me3 also mark transposons, satellite and long terminal repeats (LTRs), such as intracisternal A-particles (IAP) and early transposons (ETn) [Barski et al., 2007; Mikkelsen et al., 2007]. IAP and ETn activity at these regions in mESCs generates double-stranded RNA, which may direct silencing by targeting of H3K9me3 [Mikkelsen et al., 2007]. However, different types of repetitive sequence are again associated with different modifications and associations may vary between cell types [Martens et al., 2005; Mikkelsen et al., 2007].

In addition, distal to telomeres and centromeres, non-genic regions greater than 1 Mb (termed 'gene deserts') are significantly enriched with H3K9me2/3 but lack H4K20me3 (Fig. 1a) [Rosenfeld et al., 2009]. Notably, modifications of both active genes (H3K4me1/2/3, H3K9me1 and H4K20me1) and silenced genes (H3K27me3) are absent in gene deserts [Rosenfeld et al., 2009].

Diversity in levels of different modifications at these repressed features suggests that not only are different targeting mechanisms deployed across different regions but also that histone modifications may function differently with respect to distinct mechanisms of repression. Although ChIP-seq provides information on relative genome-wide distributions of specific histone modifications, the absolute levels of modification and thus their dynamics at the regions discussed above are unclear from this data. Furthermore, FRAP experiments demonstrate dynamic and transient interactions of HP1 with heterochromatin [Cheutin et al., 2003], raising the possibility that histone modifications at these regions may also function dynamically, occurring transiently rather than as static landmarks across chromosomes. If heterochromatic silencing is heritable and self-reinforcing (e.g. H3K9me3-bound HP1 recruits K9-methyltransferase Suv39h1 [Lachner et al., 2001]), H3K9me3 may exist as a more stable modification. Additional quantitative information is required to evaluate relationships between modification dynamics and potential function (see Histone Modifications as Dynamic Operational Marks in Genomic Processes Section).

HISTONE MODIFICATIONS ASSOCIATED WITH SILENT GENES

Although modifications such as H3K9me2/3, H3K27me2/3 and H4K20me3 are strongly associated with gene silencing and heterochromatin, silent genes can be distinguished from non-genic regions and active genes by a greater enrichment of H3K27me3 (Fig. 1b) [Rosenfeld et al., 2009]. They are also associated with H3K9me2/3; depleted of common active gene modifications (H3K4me3, H3K9me1, H4K20me1 and H3K36me3); and lack H4K20me3 compared to centromeric regions [Rosenfeld et al., 2009]. As discussed below, promoters of silent CpG-rich genes are often associated with H3K27me3 in conjunction with H3K4me3, while those of inactive CpG-poor genes are devoid of H3K4me3 [Mikkelsen et al., 2007]. While H3K27me3 associated with silent genes often occurs as a punctate peak at promoters, larger domains of H3K27me3 spanning several kbp and several genes have also been observed [Bernstein et al., 2006; Mikkelsen et al., 2007].

Silent imprinted loci notably deviate from this pattern (Fig. 1a): they lack H3K27me3 but are instead enriched with H3K9me3 and H4K20me3 [Delaval et al., 2007; Mikkelsen et al., 2007]. For example, at the *Igf2r* imprinted gene cluster, allele-specific expression is associated with H3K4me3, while the silent allele is marked by H3K9me3 and H4K20me3 without H3K27me3 [Delaval et al., 2007; Mikkelsen et al., 2007]. Furthermore, H3K9me3 and H4K20me3 mark pseudogenes and clustered silent genes, such as the *ZNF* repeats [Barski et al., 2007].

Comparison of repressive histone modifications between non-genic and genic regions suggests that distinct chromatin structures are associated with different modes of silencing. Whilst silencing mechanisms that act on repetitive sequences and centromeres would be expected to be similar and stable across various cell types,

multicellular development requires dynamic changes in the patterns of gene expression and silencing. Repression by H3K27me3-associated (vs. H3K9me2/3-associated) mechanisms may therefore correlate with a greater degree of plasticity of silencing and responsiveness to environmental cues during cell differentiation.

H3K4me3–H3K36me3 DOMAINS PREDICT TRANSCRIPTION UNITS

In general, non-silenced genic regions are largely devoid of H3K9me2/3, H3K27me3 and H4K20me3 (Fig. 1a) [Rosenfeld et al., 2009]. The greater complexity of modification distribution across poised and actively transcribed genes is described below in Histone Modifications Across Transcribed Genes Section. Transcription units have been reliably annotated through analysis of H3K4me3-H3K36me3 domains: punctate peaks of H3K4me3 at transcriptional start sites (TSSs) followed downstream by broader, 3′-peaking enrichments of H3K36me3 across transcribed regions (Fig. 1a–c) [Barski et al., 2007; Mikkelsen et al., 2007; Guttman et al., 2009]. Seventy percent of these domains correlate with actively transcribed genes in mESCs and approximately 70% in human cells are conserved in orthologous mouse regions [Guttman et al., 2009]. These strong correlations with known transcription units are not only useful for improving TSS annotation and the extent of primary transcripts, but may also indicate the existence of alternative TSSs and termination sites [Barski et al., 2007; Guenther et al., 2007; Mikkelsen et al., 2007]. Furthermore, intergenic K4me3-K36me3 domains that do not overlap protein-coding genes have been suggested as candidate (and in certain cases, demonstrated) transcription units for non-coding RNAs – transcripts that have been hard to predict by comparative genomics [Mikkelsen et al., 2007; Guttman et al., 2009].

Despite the observation that histone modifications at TSSs are highly similar between different cell types [Heintzman et al., 2009], a single cell-type-specific map cannot absolutely predict the entire complement of potential transcription units present in the genome, due to cell-type-specific gene expression patterns and the correlation of H3K36me3 across actively transcribed genes (Fig. 1b,c) [Barski et al., 2007; Guenther et al., 2007; Mikkelsen et al., 2007]. Furthermore, gene-specific variability in distributions of H3K36 dimethylation and trimethylation downstream of H3K4me3 peaks have been observed, the function of which is currently unclear [Lian et al., 2008].

H3K4 METHYLATION AND H3K27me3 IDENTIFY DIFFERENT CLASSES OF PROMOTER

The histone modification most frequently attributed to promoter regions is H3K4me3, which generally occurs as a punctate peak, within 1 kb, immediately downstream of TSSs (Fig. 1a–c) [Bernstein et al., 2005; Barski et al., 2007; Guenther et al., 2007; Heintzman et al., 2007; Koch et al., 2007]. Lower and more broadly distributed signals of H3K4me2 are also associated with TSSs and often, though not always, co-associated with H3K4me3 (Fig. 1c) [Bernstein et al., 2005; Koch et al., 2007; Orford et al., 2008]. Both H3K4me2 and H3K4me3 are highly correlated with H3K9ac in both murine and human ESCs and differentiated cells [Bernstein et al., 2005]. Levels of H3K4me3 positively correlate with gene expression: with increasing activity, H3K4me3 signals become more sharply defined as an asymmetric double peak, with increasing enrichment downstream of the promoter; greater depletion occurs directly over the TSS, which may result from nucleosome depletion or RNA polymerase II (Pol II) binding [Barski et al., 2007; Heintzman et al., 2007; Koch et al., 2007]. A potential link between this bimodal distribution and divergent initiation from human promoters has recently been proposed [Core et al., 2008; Seila et al., 2008].

H3K4me3 is strongly correlated with the presence of Pol II: over 90% of regions bound by Pol II correlate with H3K4me3 [Barski et al., 2007; Guenther et al., 2007]. Genes with H3K4me3/K9ac-marked promoters that associate with initiation-competent Pol II, however, often do not produce full-length transcripts and their coding regions are not associated with elongation-dependent modifications like H3K36me3, suggesting widespread transcriptional control by post-initiation regulatory mechanisms [Guenther et al., 2007]. This is consistent with the observation that H3K4me3 and H3K9ac occur at poised immediate-early genes independently of transcription [Hazzalin and Mahadevan, 2005; Edmunds et al., 2008] and that almost 60% of silent promoters contain H3K4me3 [Barski et al., 2007; Guenther et al., 2007].

Four broad classes of promoter can be identified in ESCs and differentiated cells that are marked differently by H3K4me2/3 and H3K27me3, correlating with underlying DNA sequence and linked gene activity (Fig. 1b,c). Firstly, CpG-rich promoters are strongly associated with H3K4me3: 95% of TSSs with H3K4me3 have CpG islands; 91% of TSSs with CpG islands have H3K4me3 [Bernstein et al., 2006]. Promoters exclusively modified by H3K4me3, in the absence of H3K27me3, positively correlate with gene expression and are overrepresented by housekeeping genes [Mikkelsen et al., 2007]. Secondly, a subset of CpG-rich promoters, represented by genes encoding developmental regulators, strongly associate with both H3K4me3 and H3K27me3 and correlate with gene repression [Mikkelsen et al., 2007]. This co-localisation has been termed 'bivalent' modification (Fig. 1a) and will be discussed further in Changes in Histone Modification during Differentiation: Epigenetic Landmarks Versus Dynamic Modification Section. Thirdly, CpG-poor promoters are largely devoid of bivalent modification, and in ESCs, over 90% of CpG-poor promoters are unmarked by either H3K4me3 or H3K27me3, correlating with gene repression [Mikkelsen et al., 2007]. These promoters are overrepresented by tissue-specific genes, and increases in gene expression that occur during lineage specification correlate strongly with increases in H3K4me3 [Mikkelsen et al., 2007]. Fourthly, although promoters are rarely marked with H3K4me2 in the absence of H3K4me3 across different human cell lines [Koch et al., 2007], a study in multipotent haematopoietic stem cells showed that H3K4me2^{+}/me3^{-} modification may occur at subsets of CpG-poor promoters, marking them poised for activation or repression depending on lineage-specification into restricted cell types [Orford et al., 2008].

HISTONE MODIFICATIONS AT REGULATORY ELEMENTS: ENHANCERS AND INSULATORS

Enhancers are defined as orientation and position-independent genomic elements that enhance the basal level of transcription from a promoter. Frequently located several kb distal to TSSs, enhancers

are difficult to identify by comparative genomics due to low sequence conservation and a lack of correlation between evolutionary conservation and enhancer activity. While they correlate with DNase I hypersensitivity and the binding of coregulators such as p300/CBP lysine acetyltransferases (KATs), both these properties may also be seen at certain promoters. Whether promoters and enhancers can be differentiated based on histone modifications is currently contentious due to conflicting ChIP-chip and ChIP-seq reports [Barski et al., 2007; Heintzman et al., 2007, 2009].

ChIP-chip analysis initially showed that promoters and enhancers share variable levels of H3K4me2, H3K9ac and H4 acetylation, but enhancers are specifically associated with H3K4me1 enrichment and H3K4me3 depletion [Heintzman et al., 2007]. These results are consistent with an independent ChIP-chip analysis that identified strong signals for H3K4me1 at TSS-distal regions with intermediate H3K4me2 levels and substantial depletion of H3K4me3 [Koch et al., 2007]. Furthermore, while H3K4me1-marked enhancers are frequently associated with H3K18ac and H3K27ac, H3K4me3-enriched promoters are acetylated preferentially at H3K9ac in addition to H3K18ac and H3K27ac [Heintzman et al., 2009].

Based on these chromatin signatures, more than 36,000 enhancers have been predicted in HeLa cells, the majority of which are also associated with DNase I hypersensitivity, transcription factor-binding motifs, and p300-enrichment [Heintzman et al., 2009]. Unexpectedly, while histone modifications at promoters are highly similar across cell types, enhancer modification profiles are highly cell-type specific despite similar global gene expression patterns [Heintzman et al., 2009]. Results of these studies conflict, however, with an independent ChIP-seq study demonstrating that both H3K4me1 and H3K4me3, along with H3K9me1, are enriched at enhancers [Barski et al., 2007]. Although further analysis has suggested that many different combinations of acetylation and methylation may occur at enhancers (unexpectedly including H3K9me2/3, H3K27me1/2/3 and H4K20me3), the most frequently observed patterns contained H3K4me1/2/3, K9me1 and K18ac [Wang et al., 2008]. No significant correlation was observed between different enhancer histone modification patterns and the level of expression associated with the nearest TSS [Wang et al., 2008]. These reported differences in H3K4me1/3 distribution at enhancers may be attributed to different criteria used in initial prediction algorithms (p300-binding sites vs. DNase I hypersensitive sites, distal to TSSs and depleted in CTCF) or due to different thresholds used to identify statistically significant peaks from raw ChIP-chip or ChIP-seq data [Barski et al., 2007; Heintzman et al., 2007, 2009; Wang et al., 2008].

Insulators are regulatory elements that through binding CTCF protein not only delineate domains within which enhancers control gene expression, but also restrict spreading of heterochromatin into euchromatic regions [Gaszner and Felsenfeld, 2006]. Alignment of CTCF sites distal to Pol II binding sites with ChIP-seq-derived methylation profiles show insulators are enriched with H3K4me1/2/3 and H3K9me1 (Fig. 1b), but are depleted in H3K9me2/3 [Barski et al., 2007]. Consistent with insulator function, punctate CTCF sites separate H3K27me1-enriched active regions from silent domains containing H3K27me3 (Fig. 1b) [Barski et al., 2007].

HISTONE MODIFICATIONS ACROSS TRANSCRIBED GENES

ChIP-seq analysis of 18 acetylation and 19 methylation modifications in human $CD4^+$ T-cells reveals a plethora of combinatorial patterns associated with actively transcribed genes [Barski et al., 2007; Wang et al., 2008]. H3K27me3 is prominently absent (Fig. 1c) and acetylation is consistently linked to gene expression, but no single site-specific modification correlates exclusively with increasing gene activation [Wang et al., 2008]. However, a modification 'backbone' pattern consisting of H2BK5ac, H2BK12ac, H2BK20ac, H2BK120ac, H3K4me1/2/3, H3K9ac, H3K9me1, H3K18ac, H3K27ac, H3K36ac and H4K8ac is detected with specific antibodies at 25% of promoters analysed [Wang et al., 2008]. Some of these (e.g. H3K9ac and H3K9me1; or H3K4me1, 2 and 3) clearly cannot coexist at the same site on the H3 tail, implying heterogeneity of nucleosomes within this combinatorial pattern, despite claims of correlations at an 'individual nucleosome level' [Wang et al., 2008]. In fact, the data is open to multiple interpretations (shown in Fig. 1d) and does not prove physical coexistence of these modifications on the same histone tail, nucleosome, chromosome or individual cell. Four categories of combination can be identified based on classifying the most frequent patterns by expression level: (i) low expression associated with H3K27me3 and absence of acetylation; (ii) low expression associated with H3K4me3 alone; (iii) intermediate expression associated with H3K36me3, or the aforementioned modification backbone (+/− H4K16ac); and (iv) most highly expressed genes associated with the backbone plus H4K16ac, H4K20me1, H3K79me1/2/3 and H2BK5me1 [Wang et al., 2008].

In addition to combinatorial diversity, distinct profiles of enrichment across promoter and transcribed regions are apparent for different modifications [Barski et al., 2007; Wang et al., 2008]. As shown in Figure 1c, H2BK5me1, H3K4me1/2/3, H3K9me1, H3K27me1, H4K20me1 and H3K79me1/2/3 predominantly occur over the 5′ coding region while H3K36me3 peaks at the 3′ end [Vakoc et al., 2006; Barski et al., 2007; Steger et al., 2008; Wang et al., 2008]. Recent ChIP-chip data suggests that H3K36me3 preferentially marks exons [Kolasinska-Zwierz et al., 2009] yet more varied patterns of H3K36me2 and H3K36me3 have been observed across specific loci [Lian et al., 2008]. Promoters of actively transcribed genes are associated with increased signals of H3K4me3 either side of the TSS; increases in H3K4me1 and H3K4me2 are more broadly distributed downstream of TSSs (Fig. 1c) [Barski et al., 2007; Koch et al., 2007]. Consistent with site-specific activity and the association of different KATs with initiating and elongating forms of Pol II [Cho et al., 1998], different acetylations occur with different distributions (Fig. 1c) [Wang et al., 2008]. Like H3K4me3, acetylation generally occurs with a bimodal distribution around TSSs, probably due to nucleosome depletion or Pol II binding at TSSs. Highly expressed genes harbour more asymmetric distributions; for example H3K9ac is skewed downstream of TSSs, while H3K36ac localises further upstream—the latter probably reflective of increased K36 methylation in transcribed regions [Barski et al., 2007; Wang et al., 2008]. Curiously, transitions in lysine acetylation/methylation occur for several lysine residues (H2BK5, H3K9, H3K27 and H3K36) from promoters to transcribed regions [Wang et al., 2008].

Although H3K9me3 is widely reported as a repressive, heterochromatic modification (especially in conjunction with H4K20me3—see Distinct Sets of Modifications Demarcate Non-Genic Chromosomal Features: Analysis of Telomeres, Centromeres and Repetitive DNA Section) and poorly correlates with gene expression, ChIP-seq shows that punctate H3K9me3 signals do occur in active chromatin domains without colocalised H4K20me3 [Barski et al., 2007]. This observation is consistent with earlier reports of H3K9me3 as a mark of transcription elongation and of HP1 localised to coding regions of active genes [Vakoc et al., 2005, 2006; Kim et al., 2007].

H3K36me3 and H4K20me1 in coding regions of active genes have been shown to be dependent on transcription elongation [Edmunds et al., 2008; Vakoc et al., 2006]. Whether modifications like H2BK5me1, H3K9me1 and H3K27me1 exist prior to transcription or are deposited co- or post-transcriptionally is currently unclear. Composite profiles derived from hundreds of genes classified by expression patterns reveal common trends and suggest specific targeting of enzymatic activities to different regions [Barski et al., 2007; Wang et al., 2008], but such analysis averages out gene-specific quantitative differences. Although the precise functions of many histone modifications within transcribed regions are still unclear, there are indications that they may be linked to cotranscriptional processing of the primary transcript; for example H3K4me3 is required for CHD1 recruitment and efficient splicing [Sims et al., 2007]; loss of H3K36me3 is associated with nuclear export defects [Yoh et al., 2008].

HISTONE MODIFICATIONS AS DYNAMIC OPERATIONAL MARKS IN GENOMIC PROCESSES

The distribution of histone modifications across mammalian genomes revealed by ChIP-seq/chip/SAGE studies are often interpreted as signatures of specific chromosomal domains and function; for example H3K4me3-K36me3 as a transcription unit or K9me3-K20me3 as pericentromeric [Barski et al., 2007; Mikkelsen et al., 2007; Guttman et al., 2009; Rosenfeld et al., 2009]. However, the overwhelming complexity of combinatorial modification [Wang et al., 2008] and quantitative, high resolution studies at individual mammalian loci [Edmunds et al., 2008] suggest that rather than stable signatures, certain modifications may act dynamically and transiently during specific genomic processes.

IMMEDIATE-EARLY GENE INDUCTION HIGHLIGHTS THE DYNAMICS AND MICROHETEROGENEITY OF HISTONE MODIFICATION

Mammalian immediate-early (IE) genes, such as c-*jun* and c-*fos*, are rapidly and transiently induced via activation of mitogen-activated protein kinase (MAPK) cascades in quiescent cells in response to specific mitogens and stress stimuli [Hazzalin and Mahadevan, 2002]. Serum-starved mouse fibroblast cultures provide homogeneous, synchronised cell populations to analyse histone modifications during IE gene induction. Early studies by Vincent Allfrey and colleagues on c-*fos* induction proposed that hyperacetylation correlates with gene induction [Allegra et al., 1987; Chen and Allfrey, 1987]. The existence of dynamic acetylation and microheterogeneity at IE genes uncouples transcription from stable histone modification patterns and thus necessitates revision of this model.

Firstly, dynamic histone acetylation is strikingly observed on a global level: acetic acid-urea gel electrophoresis demonstrates that virtually all K4-methylated histone H3 undergoes rapid, complete hyperacetylation upon histone deacetylase (HDAC) inhibition by trichostatin A (TSA), suggesting that acetyl groups are continuously turned over on K4-methylated H3 by the combined action of KATs and HDACs [Hazzalin and Mahadevan, 2005]. Consistent with this observation, H3K9ac and H3K4me3 have comparable distributions across promoters of c-*fos* and c-*jun* [Edmunds et al., 2008] and a greater proportion of total promoter sequences are recovered in H3K9ac ChIPs after TSA treatment, even in the absence of signalling to and transcription of these genes [Hazzalin and Mahadevan, 2005]. Secondly, changes in H3K9ac, H3K4me3 and H3K36me3 correlate with kinetics of IE gene expression: analysis of an amplicon at 444 bp downstream of the TSS of c-*fos* shows that in quiesced cells 10% of these sequences are associated with H3K9ac; this proportion increases to 30% after 30 min of mitogen stimulation; but subsequently drops to 15% after a further 30 min [Edmunds et al., 2008]. Since anti-H3K9ac antibodies utilised in these studies produce virtually complete immunodepletion of H3K9ac from input chromatin, changes in % input reflect changes in absolute levels of modification (rather than changes in epitope accessibility) due to dynamic shifts in targeted KAT and HDAC activity (Fig. 2A). Thirdly, although the magnitude and duration of MAPK-induced H3S10 phosphorylation correlates with IE gene induction, recovery of DNA fragments associated with H3S10ph or H3K9acS10ph are very low (<5% input) [Thomson et al., 2001; AL Clayton and LCM, unpublished work]. Given global signalling responses and quantitative immunoprecipitation [Thomson et al., 1999; Dyson et al., 2005], such low recoveries indicate that only a small proportion of nucleosomes at a given position carry H3S10ph/H3K9acS10ph. Thus mitogen-inducible H3S10 phosphorylation is a dynamic, rather than stable, modification associated with gene induction which may occur transiently: either turned over continuously or targeted asynchronously as a single pulse to individual genes (Fig. 2B).

Studies of these genes have also demonstrated the existence of microheterogeneity of histone modification: within homogeneous, synchronised cell populations, distinct populations of nucleosomes with different modifications simultaneously exist at equivalent positions (Fig. 3). Firstly, ChIPs with antibodies specifically recognising either H3K9ac or H3K9acS10ph demonstrate that both phosphoacetylated, and acetylated (but non-phosphorylated) chromatin fragments simultaneously exist within a cell population at equivalent positions at c-*jun* [Thomson et al., 2001]. Secondly, quantitative differences in H3K4me3, H3K36me3 and H3K79me2 at equivalent positions at active IE genes indicate simultaneous occurrence of all three modifications on equivalent nucleosomes across an entire cell population is unlikely [Edmunds et al., 2008]. Overlapping modification profiles can be interpreted in three distinct ways as shown in Figure 3: (i) mutually exclusive populations (Fig. 3B); (ii) partially overlapping (Fig. 3C); (iii) completely overlapping (Fig. 3D). Note that microheterogeneity still exists in (iii) when a substantial proportion of sequences are not

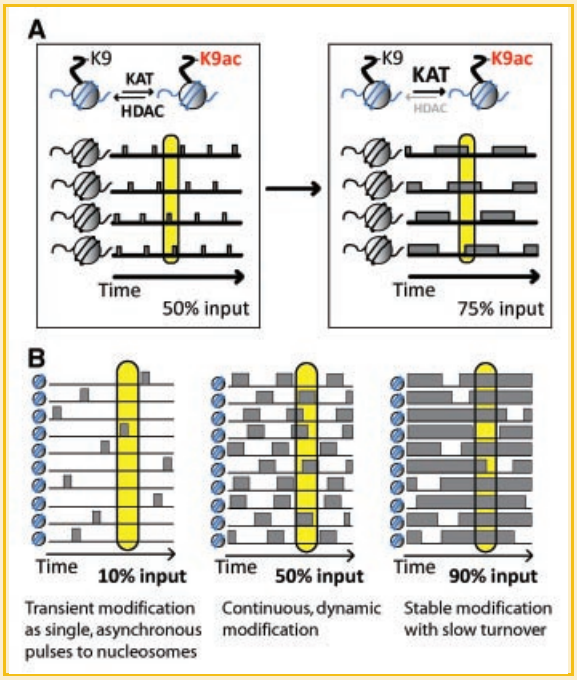

Fig. 2. Dynamic histone modification. A: Changes in lysine acetyltransferase (KAT) and histone deacetylase (HDAC) activity lead to differences in stability of H3K9ac (represented by the width of the grey rectangles). Differences in stability affect the proportion of nucleosomes at aligned equivalent positions (% input) carrying H3K9ac that are detected during the snapshot of chromatin immunoprecipitation (ChIP) analysis (yellow rectangle). B: Different % input recoveries may result from different dynamics of histone modification.

associated with either modification (e.g. 40% input recovery implies 60% unmodified in Fig. 3D). Thirdly, since K4-methylation exists on a histone H3 population distinct in its TSA-sensitivity phenotype from H3 molecules methylated at other sites (e.g. K36 and K79) [Hazzalin and Mahadevan, 2005; JW Edmunds and LCM, unpublished work], K4-methylation that is simultaneously detected by ChIP at the same position as H3K36me3 or H3K79me2 must occur on distinct histone H3 molecules. As a final consideration, the co-detection of mutually exclusive modifications (such as H3K9ac or K9me1; or different degrees of site-specific methylation: H3K4me1, me2 or me3—as is often observed in genome-wide data) at equivalent positions indicates multiple populations of differentially modified histones at that position.

Microheterogeneity may result from: (i) continuous turnover, with different kinetics affecting % input recoveries (Fig. 2A,B); (ii) transient deposition and removal, for example with polymerase passage (Fig. 2B); (iii) heterogeneous cell populations (Fig. 1D). Quantitative analysis of the proportion of nucleosomes carrying a modification at a given position in a homogeneous, synchronised cell population is important to assess modification dynamics and function (Fig. 2). For example, greater % input recoveries would be expected for stable modifications with slow turnover – the kinetics expected if modifications function as heritable marks. Conversely, smaller % input recoveries may indicate dynamic or transient modification – more likely functioning in short-term processes, rather than epigenetic inheritance across cell divisions.

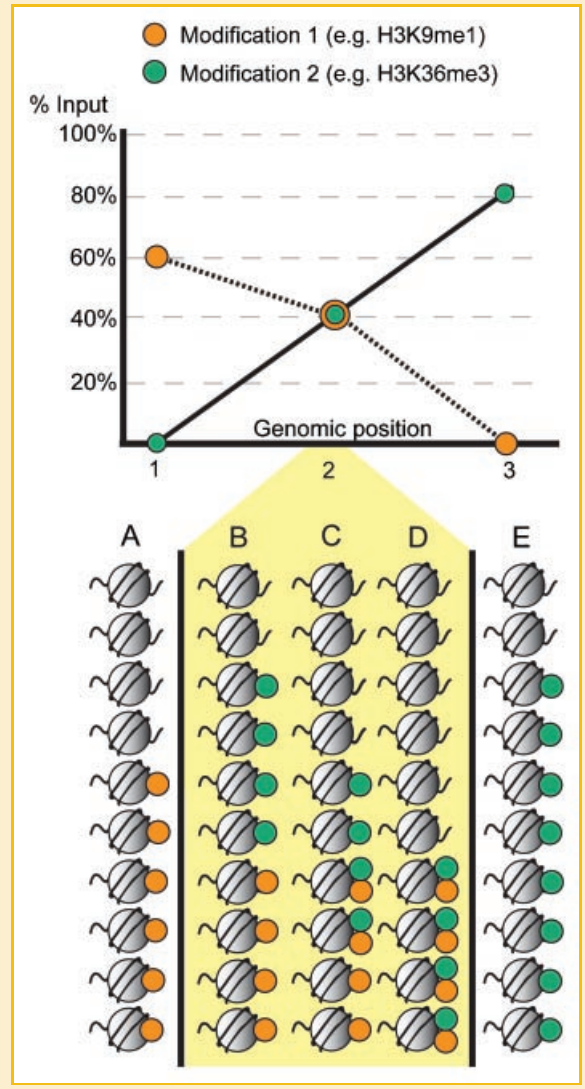

Fig. 3. Microheterogeneity of two different histone modifications. % Input (*y*-axis) is a quantitative measure of the proportion of nucleosomes within a cell population that carry a modification at given equivalent genomic positions (*x*-axis). (A) 60% of DNA sequences at position 1 are associated with modification 1. (B–D): Both modifications are associated with 40% of DNA sequences at position 2. This data can be derived from three different scenarios: (B) Mutually exclusive populations; (C) partially overlapping populations; (D) completely overlapping populations. (E): 80% of DNA sequences at position 3 are associated with modification 2.

CHANGES IN HISTONE MODIFICATION DURING DIFFERENTIATION: EPIGENETIC LANDMARKS VERSUS DYNAMIC MODIFICATION

During mammalian development, toti- and pluripotent stem cells progressively differentiate into all the cells that constitute adult organisms. Appropriate patterns of gene expression must therefore be directed and modulated by specific developmental signals, and maintained in subsequent divisions. Numerous studies have shown that mammalian genomes undergo changes in histone modifications during differentiation that reflect changes in developmental potential and gene expression; for example (i) cathepsin-mediated H3 clipping [Duncan et al., 2008]; (ii) H3K56ac redistribution

[Xie et al., 2009]; (iii) differential H3K4 di/tri-methylation at tissue-specific genes during haematopoiesis [Orford et al., 2008]; (iv) establishment of broad, differentiation-specific H3K9me2 regions associated with silencing [Wen et al., 2009]; and (v) modulation of bivalent domains [Bernstein et al., 2006; Mikkelsen et al., 2007].

Bivalent domains were initially described as broad regions of H3K27me3 containing punctate H3K4me3 peaks in mESCs [Azuara et al., 2006; Bernstein et al., 2006]; similar overlapping peaks of K4me3 and K27me3 are observed within 1 kb of TSSs of genes over-represented by developmental regulators [Barski et al., 2007; Bernstein et al., 2006; Mikkelsen et al., 2007; Pan et al., 2007; Zhao et al., 2007]. Such genes lack downstream H3K36me3 and are transcribed at very low levels (if at all) [Bernstein et al., 2006; Barski et al., 2007; Mikkelsen et al., 2007]. During ESC differentiation, bivalent promoters may persist or resolve to different chromatin states: (i) marked with H3K4me3 alone; (ii) marked with H3K27me3 alone; (iii) marked with neither H3K4me3 nor H3K27me3. Genes that resolve to H3K4me3 alone become transcriptionally active; differential lineage-specification correlates with different outcomes of bivalent resolution and changes in gene expression [Bernstein et al., 2006; Mikkelsen et al., 2007]. Association of H3K9me1, H4K20me1 and H3K4me1 with bivalent promoters in multipotent haematopoietic stem cells correlates with K27me3 loss, K4me3 retention and gene activation during subsequent differentiation [Cui et al., 2009]. Bivalent modification was thus originally proposed to facilitate stem cell plasticity: repressing lineage-specific genes to maintain pluripotency, but keeping them in a poised state for rapid activation in response to developmental signals [Bernstein et al., 2006]. However, bivalent modification is not an absolute requirement for such a process, since repressed genes carrying neither H3K4me3 nor H3K27me3 can be rapidly induced during differentiation [Pan et al., 2007]. Furthermore, bivalent promoters have been observed in differentiated cells [Barski et al., 2007; Cui et al., 2009], arguing against a stem cell-specific role. Curiously, genes required for pluripotency (*OCT4*, *NANOG* and *SOX2*) become bivalent and silenced during differentiation, but why these genes would need to be rapidly reactivated in differentiated cell-types is unclear [Pan et al., 2007].

The actual physical identity of chromatin that corresponds to bivalent regions identified by ChIP-chip/seq data is contentious, however, since it is unclear whether the modifications occur on: (i) the same histone; (ii) different histones within the same nucleosome; (iii) different nucleosomes on the same chromosome; (iv) different chromosomes; or (v) different cell populations (Fig. 1D). Sequential ChIPs indicate that only a small proportion of K4me3 and K27me3 coexist on the same histone tail [Bernstein et al., 2006]; non-overlapping peaks at higher resolution suggest localisation on different nucleosomes [Barski et al., 2007; Pan et al., 2007; Zhao et al., 2007]. Quantitative information on the proportion of nucleosomes at a given position that carry H3K4me3 or H3K27me3 is critical to evaluate hypotheses regarding the nature of bivalent modification. For example, if 90% of equivalent DNA sequences of mononucleosomal length associate with K4me3 and the same proportion is recovered in K27me3 ChIP, then this would suggest that both modifications must substantially occur simultaneously on the same nucleosomes within a cell population. Alternatively, if 90% of equivalent DNA sequences associate with K27me3 but only 10% with K4me3, then it is impossible for both these modifications to coexist in the same cells (let alone on the same histone tail or nucleosome) simultaneously across an entire cell population. Such quantitative differences would indicate microheterogeneity and raise critical questions about the precise nature of what is described as a bivalent domain.

Unfortunately, information about absolute levels of modification cannot be derived from genome-wide ChIP-chip and ChIP-seq data, precluding rigorous evaluation of the physical chromatin structures from which overlapping ChIP-chip/seq peaks are derived. Accurate quantitation is further hindered by heterogeneous or asynchronous cell populations; for example, ESC cultures are prone to spontaneous differentiation and may also contain subpopulations due to stochastic fluctuations in gene expression patterns [Graf and Stadtfeld, 2008]. This makes it hard to assess whether a peak of modification is representative of a cell population or contributed predominantly from a minority of cells with distinct characteristics.

DYNAMIC AND MICROHETEROGENEOUS MODIFICATION ACROSS MAMMALIAN GENOMES?

The microheterogeneous profiles of mammalian IE genes, typified by c-*fos* and c-*jun* suggest that a subset of histone modification occurs dynamically. How far these concepts can be applied to the rest of the genome is unclear due to a lack of quantitative, comparative data. Assessment of microheterogeneity and modification dynamics requires quantitative analysis of absolute proportions of modified nucleosomes by ChIP analysis of homogeneous, synchronised cell populations using specific antibodies with high immunoprecipitation efficiency. ChIP-seq and ChIP-chip data currently indicates relative modification distribution across the genome (e.g. H4K20me3 is eightfold more greatly enriched at pericentromeric regions compared to the genomic average [Rosenfeld et al., 2009]). However, the absolute level of modification is unclear—whether 8% of nucleosomes at a given pericentromeric position carry H4K20me3 (compared to 1% of average genomic nucleosomes), or 80% of nucleosomes at a pericentromeric locus (compared to 10% of average genomic nucleosomes). Information on the absolute level of modification is important in assessing the stability and prevalence of a modification and thus its role (Fig. 2). Stable modifications present at every copy of an equivalent locus across a cell population would be expected to return high % input recoveries; epigenetic roles for such modifications are more plausible. Stably, silenced constitutive heterochromatin may be marked in this way to facilitate perpetuation of repression across cell divisions. Lower sequences recoveries, however, indicate that only a proportion of nucleosomes at a given position carry the modification of interest. If homogeneous cell populations are analysed by quantitative immunoprecipitation, such low % input recoveries suggest the existence of dynamic modifications: either with a single transient occurrence or continuous turnover—unlikely to act in a heritable, epigenetic fashion, but as operational marks supporting short-term processes in the genome.

CONCLUSION

Technological advances have tremendously enhanced our understanding of the distributions of histone modifications across the mammalian genome in different cell types, and have facilitated annotation of features hard to identify by comparative genomics. Comparative maps under different conditions, integrated with the wealth of information generated from other genomic data sets (such as gene expression analyses, transcription factor binding sites and nuclease accessibility studies) provide an immense resource to evaluate current models, and formulate novel hypotheses relating the interaction of DNA sequence, chromatin modification, genomic processes and cellular identity. Quantitative comparison of absolute levels of histone modification, in terms of the precise proportion of nucleosomes modified at specific positions, is now necessary to validate the various proposed roles of histone modifications in mammalian genomes.

ACKNOWLEDGMENTS

BML is supported by a Wellcome Trust PhD Studentship.

REFERENCES

Allegra P, Sterner R, Clayton DF, Allfrey VG. 1987. Affinity chromatographic purification of nucleosomes containing transcriptionally active DNA sequences. J Mol Biol 196:379–388.

Allfrey VG, Faulkner R, Mirsky AE. 1964. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. Proc Natl Acad Sci USA 51:786–794.

Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merkenschlager M, Fisher AG. 2006. Chromatin signatures of pluripotent cell lines. Nat Cell Biol 8:532–538.

Barski A, Zhao K. 2009. Genomic location analysis by ChIP-Seq. J Cell Biochem 107:11–18.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. Cell 129:823–837.

Bernstein BE, Humphrey EL, Erlich RL, Schneider R, Bouman P, Liu JS, Kouzarides T, Schreiber SL. 2002. Methylation of histone H3 Lys 4 in coding regions of active genes. Proc Natl Acad Sci USA 99:8695–8700.

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ III, Gingeras TR, Schreiber SL, Lander ES. 2005. Genomic maps and comparative analysis of histone modifications in human and mouse. Cell 120:169–181.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125:315–326.

Bird A. 2007. Perceptions of epigenetics. Nature 447:396–398.

Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, Bell GW, Otte AP, Vidal M, Gifford DK, Young RA, Jaenisch R. 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 441:349–353.

Buck MJ, Lieb JD. 2004. ChIP-chip: Considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. Genomics 83:349–360.

Chen TA, Allfrey VG. 1987. Rapid and reversible changes in nucleosome structure accompany the activation, repression, and superinduction of murine fibroblast protooncogenes c-fos and c-myc. Proc Natl Acad Sci USA 84:5252–5256.

Cheutin T, McNairn AJ, Jenuwein T, Gilbert DM, Singh PB, Misteli T. 2003. Maintenance of stable heterochromatin domains by dynamic HP1 binding. Science 299:721–725.

Cho H, Orphanides G, Sun X, Yang XJ, Ogryzko V, Lees E, Nakatani Y, Reinberg D. 1998. A human RNA polymerase II complex containing factors that modify chromatin structure. Mol Cell Biol 18:5355–5363.

Clayton AL, Hazzalin CA, Mahadevan LC. 2006. Enhanced histone acetylation and transcription: A dynamic perspective. Mol Cell 23:289–296.

Core LJ, Waterfall JJ, Lis JT. 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science 322:1845–1848.

Cui K, Zang C, Roh TY, Schones DE, Childs RW, Peng W, Zhao K. 2009. Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. Cell Stem Cell 4:80–93.

Delaval K, Govin J, Cerqueira F, Rousseaux S, Khochbin S, Feil R. 2007. Differential histone modifications mark mouse imprinting control regions during spermatogenesis. EMBO J 26:720–729.

Duncan EM, Muratore-Schroeder TL, Cook RG, Garcia BA, Shabanowitz J, Hunt DF, Allis CD. 2008. Cathepsin L proteolytically processes histone H3 during mouse embryonic stem cell differentiation. Cell 135:284–294.

Dyson MH, Thomson S, Inagaki M, Goto H, Arthur SJ, Nightingale K, Iborra FJ, Mahadevan LC. 2005. MAP kinase-mediated phosphorylation of distinct pools of histone H3 at S10 or S28 via mitogen- and stress-activated kinase 1/2. J Cell Sci 118:2247–2259.

Edmondson DG, Davie JK, Zhou J, Mirnikjoo B, Tatchell K, Dent SY. 2002. Site-specific loss of acetylation upon phosphorylation of histone H3. J Biol Chem 277:29496–29502.

Edmunds JW, Mahadevan LC, Clayton AL. 2008. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. EMBO J 27:406–420.

Gaszner M, Felsenfeld G. 2006. Insulators: Exploiting transcriptional and epigenetic mechanisms. Nat Rev Genet 7:703–713.

Graf T, Stadtfeld M. 2008. Heterogeneity of embryonic and adult stem cells. Cell Stem Cell 3:480–483.

Guccione E, Bassi C, Casadio F, Martinato F, Cesaroni M, Schuchlautz H, Luscher B, Amati B. 2007. Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. Nature 449:933–937.

Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. 2007. A chromatin landmark and transcription initiation at most promoters in human cells. Cell 130:77–88.

Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL, Lander ES. 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458:223–227.

Hazzalin CA, Mahadevan LC. 2002. MAPK-regulated transcription: A continuously variable gene switch? Nat Rev Mol Cell Biol 3:30–40.

Hazzalin CA, Mahadevan LC. 2005. Dynamic acetylation of all lysine 4-methylated histone H3 in the mouse nucleus: Analysis at c-fos and c-jun. PLoS Biol 3:e393.

Hebbes TR, Thorne AW, Crane-Robinson C. 1988. A direct link between core histone acetylation and transcriptionally active chromatin. EMBO J 7:1395–1402.

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet 39:311–318.

Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, Ching KA, Antosiewicz-Bourget JE, Liu H, Zhang X, Green RD, Lobanenkov VV, Stewart R, Thomson JA, Crawford GE, Kellis M, Ren B. 2009. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 459:108–112.

Iberg AN, Espejo A, Cheng D, Kim D, Michaud-Levesque J, Richard S, Bedford MT. 2008. Arginine methylation of the histone H3 tail impedes effector binding. J Biol Chem 283:3006–3010.

Kim A, Kiefer CM, Dean A. 2007. Distinctive signatures of histone methylation in transcribed coding and noncoding human beta-globin sequences. Mol Cell Biol 27:1271–1279.

Klose RJ, Bird AP. 2006. Genomic DNA methylation: The mark and its mediators. Trends Biochem Sci 31:89–97.

Koch CM, Andrews RM, Flicek P, Dillon SC, Karaoz U, Clelland GK, Wilcox S, Beare DM, Fowler JC, Couttet P, James KD, Lefebvre GC, Bruce AW, Dovey OM, Ellis PD, Dhami P, Langford CF, Weng Z, Birney E, Carter NP, Vetrie D, Dunham I. 2007. The landscape of histone modifications across 1% of the human genome in five human cell lines. Genome Res 17:691–707.

Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J. 2009. Differential chromatin marking of introns and expressed exons by H3K36me3. Nat Genet 41:376–381.

Kouzarides T. 2007. Chromatin modifications and their function. Cell 128: 693–705.

Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS, Presser A, Nusbaum C, Xie X, Chi AS, Adli M, Kasif S, Ptaszek LM, Cowan CA, Lander ES, Koseki H, Bernstein BE. 2008. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet 4:e1000242.

Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 410:116–120.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, et al. 2001. Initial sequencing and analysis of the human genome. Nature 409: 860–921.

Lian Z, Karpikov A, Lian J, Mahajan MC, Hartman S, Gerstein M, Snyder M, Weissman SM. 2008. A genomic analysis of RNA polymerase II modification and chromatin architecture related to 3′ end RNA polyadenylation. Genome Res 18:1224–1237.

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature 389:251–260.

Martens JH, O'Sullivan RJ, Braunschweig U, Opravil S, Radolf M, Steinlein P, Jenuwein T. 2005. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. EMBO J 24:800–812.

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448:553–560.

Orford K, Kharchenko P, Lai W, Dao MC, Worhunsky DJ, Ferro A, Janzen V, Park PJ, Scadden DT. 2008. Differential H3K4 methylation identifies developmentally poised hematopoietic genes. Dev Cell 14:798–809.

Pan G, Tian S, Nie J, Yang C, Ruotti V, Wei H, Jonsdottir GA, Stewart R, Thomson JA. 2007. Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. Cell Stem Cell 1:299–312.

Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E, Zeitlinger J, Lewitter F, Gifford DK, Young RA. 2005. Genome-wide map of nucleosome acetylation and methylation in yeast. Cell 122:517–527.

Roh TY, Cuddapah S, Zhao K. 2005. Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. Genes Dev 19:542–552.

Roh TY, Cuddapah S, Cui K, Zhao K. 2006. The genomic landscape of histone modifications in human T cells. Proc Natl Acad Sci USA 103:15782–15787.

Rosenfeld JA, Wang Z, Schones DE, Zhao K, Desalle R, Zhang MQ. 2009. Determination of enriched histone modifications in non-genic portions of the human genome. BMC Genomics 10:143.

Sarma K, Reinberg D. 2005. Histone variants meet their match. Nat Rev Mol Cell Biol 6:139–149.

Schreiber SL, Bernstein BE. 2002. Signaling network model of chromatin. Cell 111:771–778.

Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA, Young RA, Sharp PA. 2008. Divergent transcription from active promoters. Science 322:1849–1851.

Sims RJ III, Millhouse S, Chen CF, Lewis BA, Erdjument-Bromage H, Tempst P, Manley JL, Reinberg D. 2007. Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. Mol Cell 28:665–676.

Steger DJ, Lefterova MI, Ying L, Stonestrom AJ, Schupp M, Zhuo D, Vakoc AL, Kim JE, Chen J, Lazar MA, Blobel GA, Vakoc CR. 2008. DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. Mol Cell Biol 28:2825–2839.

Strahl BD, Allis CD. 2000. The language of covalent histone modifications. Nature 403:41–45.

Sullivan BA, Karpen GH. 2004. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nat Struct Mol Biol 11:1076–1083.

Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. 2007. How chromatin-binding modules interpret histone modifications: Lessons from professional pocket pickers. Nat Struct Mol Biol 14:1025–1040.

Thomson S, Clayton AL, Hazzalin CA, Rose S, Barratt MJ, Mahadevan LC. 1999. The nucleosomal response associated with immediate-early gene induction is mediated via alternative MAP kinase cascades: MSK1 as a potential histone H3/HMG-14 kinase. EMBO J 18:4779–4793.

Thomson S, Clayton AL, Mahadevan LC. 2001. Independent dynamic regulation of histone phosphorylation and acetylation during immediate-early gene induction. Mol Cell 8:1231–1241.

Turner BM. 2005. Reading signals on the nucleosome with a new nomenclature for modified histones. Nat Struct Mol Biol 12:110–112.

Vakoc CR, Mandat SA, Olenchock BA, Blobel GA. 2005. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. Mol Cell 19:381–391.

Vakoc CR, Sachdeva MM, Wang H, Blobel GA. 2006. Profile of histone lysine methylation across transcribed mammalian chromatin. Mol Cell Biol 26: 9185–9195.

Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh TY, Peng W, Zhang MQ, Zhao K. 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet 40:897–903.

Wei G, Wei L, Zhu J, Zang C, Hu-Li J, Yao Z, Cui K, Kanno Y, Roh TY, Watford WT, Schones DE, Peng W, Sun HW, Paul WE, O'Shea JJ, Zhao K. 2009. Global

mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating CD4+ T cells. Immunity 30: 155–167.

Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP. 2009. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. Nat Genet 41:246–250.

Xie W, Song C, Young NL, Sperling AS, Xu F, Sridharan R, Conway AE, Garcia BA, Plath K, Clark AT, Grunstein M. 2009. Histone H3 lysine 56 acetylation is linked to the core transcriptional network in human embryonic stem cells. Mol Cell 33:417–427.

Yoh SM, Lucas JS, Jones KA. 2008. The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. Genes Dev 22:3422–3434.

Zhao XD, Han X, Chew JL, Liu J, Chiu KP, Choo A, Orlov YL, Sung WK, Shahab A, Kuznetsov VA, Bourque G, Oh S, Ruan Y, Ng HH, Wei CL. 2007. Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell 1:286–298.

Zlatanova J, Bishop TC, Victor JM, Jackson V, van Holde K. 2009. The nucleosome family: Dynamic and growing. Structure 17:160–171.