



## Review

# Understanding the interplay between CpG island-associated gene promoters and H3K4 methylation<sup>☆</sup>



Amy L. Hughes<sup>1</sup>, Jessica R. Kelley<sup>1</sup>, Robert J. Klose<sup>\*</sup>

Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, United Kingdom

## ARTICLE INFO

**Keywords:**  
Chromatin  
Transcription  
DNA methylation  
CpG islands  
H3K4me3

## ABSTRACT

The precise regulation of gene transcription is required to establish and maintain cell type-specific gene expression programs during multicellular development. In addition to transcription factors, chromatin, and its chemical modification, play a central role in regulating gene expression. In vertebrates, DNA is pervasively methylated at CG dinucleotides, a modification that is repressive to transcription. However, approximately 70% of vertebrate gene promoters are associated with DNA elements called CpG islands (CGIs) that are refractory to DNA methylation. CGIs integrate the activity of a range of chromatin-regulating factors that can post-translationally modify histones and modulate gene expression. This is exemplified by the trimethylation of histone H3 at lysine 4 (H3K4me3), which is enriched at CGI-associated gene promoters and correlates with transcriptional activity. Through studying H3K4me3 at CGIs it has become clear that CGIs shape the distribution of H3K4me3 and, in turn, H3K4me3 influences the chromatin landscape at CGIs. Here we will discuss our understanding of the emerging relationship between CGIs, H3K4me3, and gene expression.

## 1. Introduction

The precise control of gene transcription is required for cellular homeostasis and to establish the cell type-specific gene expression patterns that are necessary for complex multicellular development. This regulation primarily relies on site-specific DNA-binding transcription factors (TFs), which shape how the transcriptional machinery engages with gene promoters to initiate transcription. However, in eukaryotes, the chromatin template on which transcription occurs can also profoundly affect how gene expression is established and maintained. For example, the intimate association between DNA and histone octamers within nucleosomes can occlude the binding of TFs and RNA Polymerase II (RNAPII) [1] (reviewed in Li et al., Levine et al. and Venkatesh and Workman [2–4]). Therefore, the position of nucleosomes at gene promoters and regulatory elements can modulate transcriptional output. Furthermore, histone proteins can undergo extensive post-translational modification, including methylation, acetylation, phosphorylation, and ubiquitylation. While these modifications occur on residues throughout histone proteins, the N- or C-terminal tails that extend from the nucleosome core tend to be the most heavily modified and are thought to be particularly important in gene regulation (reviewed in Rothbart et al. and Zhang et al. [5,6]). A subset of histone

modifications appears to directly regulate chromatin structure by influencing histone-DNA or nucleosome-nucleosome interactions. Alternatively, histone modifications can create a binding site for, or exclude the binding of, ‘reader’ proteins that in turn nucleate additional activities that influence gene expression.

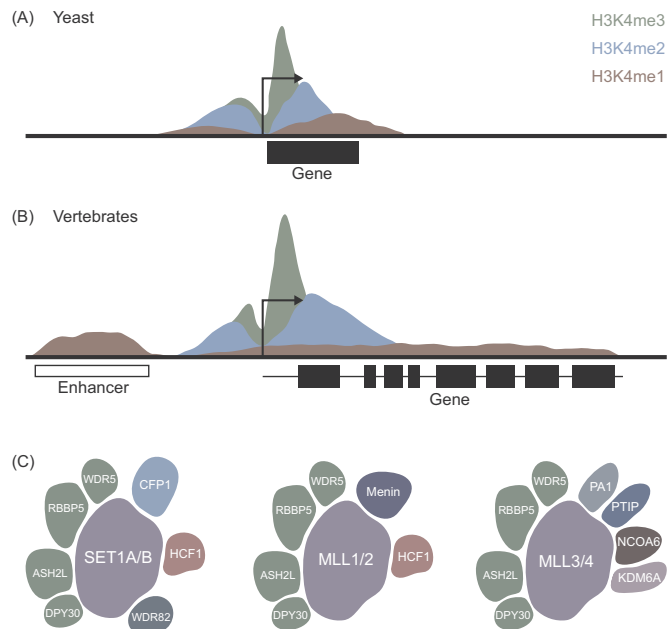
One of the most extensively studied histone modifications is the methylation of histone H3 on lysine 4 (H3K4). H3K4 can be mono-, di- or tri-methylated (H3K4me1/2/3). In simple eukaryotes, such as budding yeast, there is a gradient of H3K4 methylation across genes: H3K4me3 is found at gene promoters, H3K4me2 within gene bodies, and H3K4me1 predominates towards the 3’ end of genes (Fig. 1A) [7]. Genome-wide analysis of H3K4 methylation in multicellular eukaryotes, which have larger and more complex genomes, has identified a more distinct pattern of H3K4 methylation. While the enrichment of H3K4me3 at gene promoters is conserved across all eukaryotes, in vertebrates H3K4me1 is more broadly distributed and enriched at enhancers, while H3K4me2 tends to be elevated in regions flanking peaks of H3K4me3 (Fig. 1B) [8–11]. In line with these more diverse patterns of H3K4 methylation, in higher eukaryotes there is an expanded repertoire of enzymes that deposit these modifications. In yeast, a single histone methyltransferase (HMT) complex, known as SET1C, is responsible for placing all H3K4 methylation [12,13]. *Drosophila* have

<sup>☆</sup> This article is part of a Special Issue entitled: The MLL family of proteins in normal development and disease edited by Thomas A Milne.

<sup>\*</sup> Corresponding author.

E-mail address: [rob.klose@bioch.ox.ac.uk](mailto:rob.klose@bioch.ox.ac.uk) (R.J. Klose).

<sup>1</sup> These authors contributed equally to this work.



**Fig. 1.** Histone H3 lysine 4 methylation and the H3K4 methyltransferase complexes.

(A) and (B) Schematic depictions of the distribution of H3K4 methylation across (A) yeast and (B) vertebrate genes. Arrows indicate transcription start sites. Gene bodies are shown with black boxes representing exons.

(C) The subunit composition of mammalian H3K4 histone methyltransferases. The H3K4 HMTs are split into three groups based on the homology of their catalytic subunit: SET1A/B, MLL1/2 and MLL3/4. Shared subunits are depicted in the same colour and position.

Adapted from [266].

three H3K4 HMT complexes: the TRR complex deposits H3K4me1, while the dSET1 and TRX complexes place H3K4me2/3 at distinct genomic loci [14]. This system has further expanded in vertebrates to produce six H3K4 HMTs [15–20]. Analogous to the functional organisation of the H3K4 HMTs in *Drosophila*, mammalian MLL3/4 complexes are typically associated with H3K4me1 deposition at enhancers, while SET1A/B and MLL1/2 complexes define H3K4me2/3 at gene promoters [21–27]. These H3K4 HMT complexes share a subset of accessory factors but are distinguished based on the identity of their catalytic subunit and are further functionally specialised by the incorporation of complex-specific factors [28] (Fig. 1C and reviewed in Shilatifard [14]).

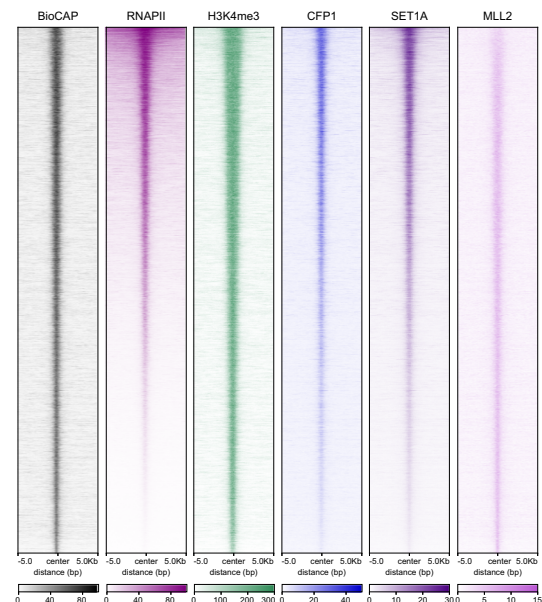
H3K4me3 is universally associated with gene promoters and has been proposed to play an important role in gene regulation by recruiting additional chromatin modifying enzymes and the transcriptional machinery, and excluding activities typically associated with transcriptional repression (reviewed in Kusch and Hyun et al. [29,30]). Therefore, significant efforts have been made to understand the mechanisms regulating its deposition and function at gene promoters. In budding yeast, H3K4 methylation patterns are predominantly defined by the recruitment of SET1C to sites of active transcription via RNAPII [31,32]. However, the diversification of H3K4 methyltransferases in higher eukaryotes has been accompanied by increasingly complex mechanisms for chromatin binding and target site recognition [33]. Transcription alone is insufficient to explain how the SET1A/B and MLL1/2 complexes associate with vertebrate gene promoters. Emerging evidence has revealed an intimate relationship between the SET1A/B and MLL1/2 complexes and a class of vertebrate-specific gene regulatory elements called CpG islands (CGIs). Here we will discuss the relationship between CGIs and H3K4me3, touching on how CGIs shape the patterns of H3K4me3 in vertebrates and how H3K4me3 itself can affect the chromatin architecture around CGI-associated gene promoters to influence gene expression. Since MLL3/4 have been primarily

implicated in placing H3K4me1 at enhancers and are not enriched at CGIs, they are beyond the scope of this review. However, their function has been covered extensively elsewhere [34,35].

## 2. CpG islands and H3K4me3

### 2.1. CpG islands delineate H3K4me3 in vertebrate genomes

Much like histones, DNA can be modified in several ways, most notably by the methylation of individual nucleotide bases [5,36,37]. The mammalian genome is pervasively methylated at the 5-position of cytosine bases in the context of CG dinucleotides, which is important for silencing parasitic DNA elements, regulating imprinted gene expression, and supporting X chromosome-inactivation [38,39] (reviewed in Smith and Meissner and Greenberg and Bourc'his [40,41]). Cytosine methylation is also mutagenic, resulting in a global depletion of CG dinucleotides from the mammalian genome [42,43]. However, short (~1–2 kb) regions of the genome, termed CpG islands, remain non-methylated and are enriched for CG dinucleotides [38,44,45]. Approximately 70% of vertebrate gene promoters are associated with CGIs, making them the most common class of promoter in vertebrates [46,47]. CGIs remain free of DNA methylation regardless of the transcriptional state of the associated gene (Fig. 2), suggesting that non-methylated DNA itself does not instruct transcription. However, it has become apparent that CGIs are associated with, and appear to delineate the profiles of, several histone modifications known to have roles in gene regulation. This is exemplified by H3K4me3 which, in vertebrate genomes, correlates most closely with the underlying boundaries of promoter-associated CGIs [47,48]. This is in contrast to simple eukaryotes where H3K4me3 is distributed in a punctate fashion around the transcription start site of actively transcribed genes [7]. Furthermore, although its levels generally scale with transcription, H3K4me3 is present at the majority of CGIs irrespective of whether the associated gene is actively transcribed (Fig. 2). This suggests that the targeting and deposition of H3K4me3 in vertebrates is intrinsically linked to CGIs and not solely reliant on the transcriptional machinery. In support of this



**Fig. 2.** CpG islands and H3K4me3.

Heatmaps illustrating non-methylated DNA (BioCAP) [47], RNAPII (transcription) [267], H3K4me3, CFP1, SET1A [61] and MLL2 [25] ChIP-seq signal across all CpG islands in mouse embryonic stem cells. The heatmaps are ranked based on RNAPII signal. H3K4me3 occurs broadly across CpG islands genome-wide. MLL2 associates with the majority of CGIs, whereas CFP1 and SET1A are enriched at CpG islands that are bound by RNAPII.

possibility, when exogenous CG-rich non-methylated DNA lacking a gene promoter is inserted into the mouse genome, it can acquire H3K4me3 in the absence of RNAPII binding [49–51].

With the exception of CGIs, the majority of the vertebrate genome is methylated in almost all somatic cell types. However, there are specific instances where the mammalian genome becomes markedly hypomethylated, for example in oocytes and at early developmental stages [40,52,53]. If recognition of non-methylated CpG DNA is important for H3K4me3 deposition, one might predict that regions of the genome that lose DNA methylation would acquire H3K4me3. Indeed, Chromatin Immunoprecipitation (ChIP)-sequencing for H3K4me3 in mouse oocytes and during early stages of zygotic development revealed the presence of large domains of H3K4me3 corresponding to non-methylated regions of the genome that are distinct from classical CGIs [54,55]. In oocytes, the appearance of these non-canonical domains of H3K4me3 is dependent on MLL2 [56,57]. However, it should be noted that these large H3K4me3 domains are likely to be cell type-specific or rely on the expression level of specific H3K4 methyltransferases, as removal of DNA methylation in embryonic stem cells (ESCs) does not cause significant spreading of H3K4me3 into non-CGI regions that become hypomethylated [48]. Overall, these observations suggest that the presence of non-methylated CpG DNA, particularly at CG-rich regions of the genome, is essential for guiding the activity of H3K4 HMTs and promoting H3K4me3 deposition.

## 2.2. SET1A/B and MLL1/2 associate with CpG islands via ZF-CxxC domains

An important step towards understanding the relationship between CGIs and H3K4me3 came with the discovery that non-methylated CpG dinucleotides are specifically recognised by proteins containing Zinc Finger (ZF)-CxxC domains [58]. The founding member of the ZF-CxxC domain-containing protein family is CFP1 (CXXC1), an essential component of the SET1A/B H3K4 HMT complexes [15,16,59,60]. Genome-wide studies mapping the location of CFP1 and SET1A *in vivo* have indicated that both proteins are enriched at CGI-associated gene promoters (Fig. 2) and *in vitro* binding studies have shown that the ZF-CxxC domain of CFP1 associates specifically with non-methylated CpG dinucleotides [25,61–63]. The DNA binding activity of CFP1 is essential for the association of SET1A with chromatin and normal levels of H3K4me3 at CGIs [61]. This suggests that the SET1A/B complexes are targeted to CGI-associated gene promoters via recognition of non-methylated CpG DNA by the ZF-CxxC domain of CFP1. This targeting logic appears to have evolved and been maintained only in the genomes of species that have pervasive DNA methylation and CGIs; CFP1 orthologues in species such as *Drosophila*, which have lost DNA methylation and classical CGIs, do not encode a functional ZF-CxxC domain (reviewed in Krauss and Reuter and Iyer et al. [64,65]).

The MLL1 and MLL2 proteins also contain ZF-CxxC domains that bind specifically to non-methylated CpG dinucleotides *in vitro* [66–70]. Like CFP1, MLL1 and MLL2 also require their ZF-CxxC domains to associate with CGIs *in vivo* [71,72]. Further evidence that the ZF-CxxC domains of MLL proteins are important for targeting to CGIs comes from the study of MLL1 in the context of leukaemia. In such instances, chromosomal translocations generate fusions between the MLL1 N-terminus, which contains its ZF-CxxC domain, and a variety of other proteins, including the transcriptional regulators AF9 and ENL [73–77]. A universal feature of these fusion proteins is the retention of the MLL1 ZF-CxxC domain, which aberrantly targets its fusion partner to CGIs, causing widespread transcriptional defects [70,73,74,78–82]. Loss or mutation of the ZF-CxxC domain in MLL-AF9 and -ENL fusion proteins reduces their capacity for inducing leukaemic cellular transformation, suggesting that the ZF-CxxC domain is crucial for their association with chromatin [69,71,78,79].

Although the presence of non-methylated DNA appears to be sufficient to create H3K4me3-modified chromatin and the ZF-CxxC domains

inherent to the SET1A/B and MLL1/2 complexes can target them to CGIs, this single interaction does not completely explain the localisation and activity of these complexes. Whilst MLL2 binds broadly to CGIs *in vivo* (Fig. 2), it only appears to contribute significantly to H3K4me3 at a subset of lowly transcribed CGI-associated gene promoters [24,25]. In contrast, SET1A/B complexes are preferentially enriched at actively transcribed CGI-associated gene promoters (Fig. 2) [25,61,83]. Furthermore, *in vitro* measurements suggest that the affinity of ZF-CxxC domains for non-methylated CpG dinucleotides is relatively low ( $\sim 0.1\text{--}7\ \mu\text{M}$ ) and is therefore unlikely to constitute the sole chromatin binding determinant for these complexes [63,67,68]. Therefore, significant efforts have been made to understand the additional mechanisms by which H3K4 HMT complexes recognise their CGI-associated target sites to deposit H3K4me3.

## 2.3. Multivalent chromatin interactions restrict H3K4 HMTs to a subset of CpG island promoters

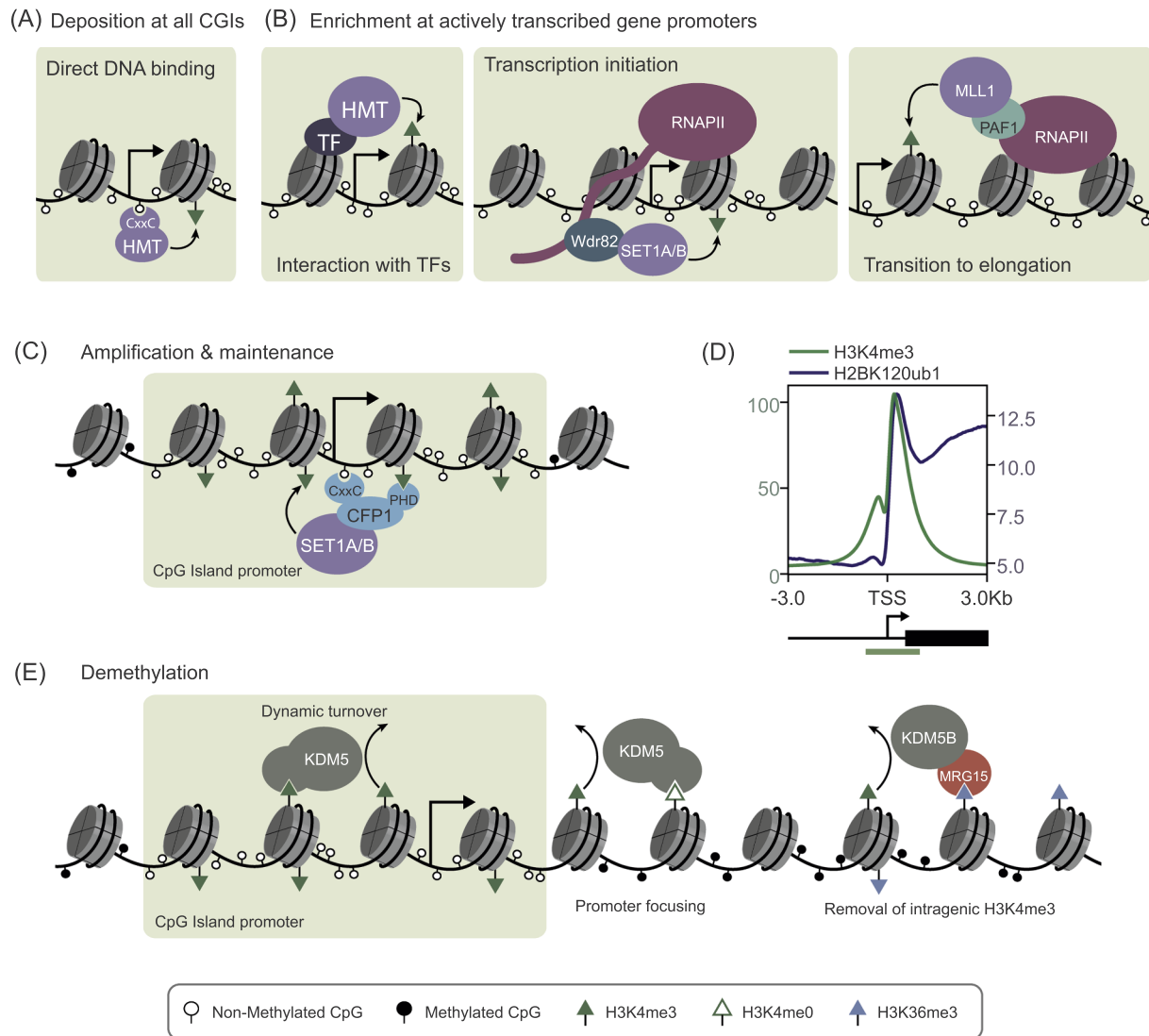
Many chromatin-associated proteins exploit multivalent interactions that rely on a series of low-affinity binding modules to provide specificity and elevated affinity for their target sites [84]. In addition to binding non-methylated CpG DNA via its ZF-CxxC domain, CFP1 also recognises H3K4me3 through its Plant HomeoDomain (PHD domain) [61,85,86]. This provides the SET1A/B complexes with a preferential affinity for CGIs that are already modified by H3K4me3. Interestingly, mutations that disrupt H3K4me3 binding by CFP1 do not completely eliminate CFP1 occupancy at CGIs. In fact, in the absence of a functional PHD domain, residual CFP1 binding is uniformly distributed across CGIs and correlates with CpG content. In contrast, if the ZF-CxxC domain of CFP1 is mutated there remains little, if any, CFP1 occupancy at CGIs [61]. This suggests that the ZF-CxxC domain of CFP1 is the primary determinant of its affinity for chromatin. Multivalent engagement via the PHD domain then stabilises binding to provide additional specificity that enriches CFP1 and SET1A/B complexes at CGIs with pre-existing H3K4me3.

MLL1 and MLL2 also contain multiple PHD domains, some of which bind to H3K4me3 [71,86–90]. The H3K4me3-binding PHD domain of MLL1 can contribute to its occupancy at some target sites, however the relevance of these PHD domains for MLL1/2 complex localisation genome-wide has not been defined [71,88]. Nevertheless, the genome-wide occupancy of MLL2 appears to be more uniformly distributed across CGIs than CFP1 and SET1A (Fig. 2) [25]. This suggests that if the PHD domains of MLL1 and MLL2 do associate with H3K4me3 *in vivo*, they do not enrich MLL1/2 complexes at the subset of CGIs that have elevated H3K4me3.

## 3. Regulating H3K4me3 at CpG island-associated gene promoters

### 3.1. Is H3K4me3 deposition directed by active transcription?

Unlike the MLL1/2 complexes, which broadly occupy all CGI promoters, the SET1A/B complexes are enriched at actively transcribed gene promoters and are proposed to contribute centrally to H3K4me3 at these loci [26,61,91–94]. Interestingly, the selectivity of the SET1A/B complexes for actively transcribed gene promoters appears to rely on multivalent interactions with CGIs and pre-existing H3K4me3 [61]. How, therefore, is H3K4me3 enriched at actively transcribed CGI-associated gene promoters to support preferential binding of the SET1A/B complexes and further elevation of H3K4me3? Whilst all four H3K4me3 HMTs can bind to CGIs via their ZF-CxxC domains (Fig. 3A) and the presence of CpG island DNA is sufficient for H3K4me3 deposition *in vivo*, this DNA-based recognition modality alone cannot distinguish between transcribed and non-transcribed genes [49–51]. Additional mechanisms must therefore exist to enrich H3K4me3 specifically at actively transcribed gene promoters. One possible explanation may be that the H3K4 HMTs are specifically targeted to gene promoters during



**Fig. 3.** Mechanisms that regulate H3K4me3 at CpG islands.

(A) H3K4me3 HMT complexes can deposit H3K4me3 across all CpG islands by binding directly to non-methylated CpG dinucleotides via their ZF-CxxC domains. The position of the transcription start site is illustrated as an arrow and a legend for the chromatin modifications is shown at the bottom of this figure.

(B) Mechanisms to enrich H3K4me3 at actively transcribed gene promoters. These include interaction of HMT complexes with transcription factors (TFs) or interactions with the transcription machinery during initiation and elongation.

(C) A mechanism to amplify and maintain H3K4me3 once initiated. CFP1 stabilises SET1A/B complexes at actively transcribed CGI-associated gene promoters through multivalent interactions with non-methylated CpG dinucleotides and pre-existing H3K4me3.

(D) A metaplot illustrating the distribution of H3K4me3 and H2BK120ub1 ChIP-seq signal at CGI associated gene promoters in mouse embryonic fibroblasts [105]. H3K4me3 is distributed around the TSS and over the CGI, whereas H2BK120ub1 peaks downstream of the TSS and is also enriched throughout the gene body. The schematic below indicates the position of the TSS as an arrow, CGI promoter as a green box and gene body as a black box.

(E) Mechanisms through which demethylation by KDM5 proteins could shape H3K4me3. These include dynamic turnover of H3K4me3, focussing of H3K4me3 at gene promoters, and removal of spuriously deposited H3K4me3 in gene bodies.

the process of gene activation. If so, deposition of H3K4me3 through transcription-coupled mechanisms could then support multivalent recognition of non-methylated CpG DNA and H3K4me3 via CFP1-containing SET1A/B complexes.

In support of this hypothesis, the SET1A/B and MLL1/2 complexes interact with a number of sequence-specific DNA binding transcription factors and long non-coding RNAs (Fig. 3B) (reviewed in Bochyńska et al. and Crump and Milne [33,34]), although the relevance of these interactions for H3K4me3 deposition genome-wide have not been fully dissected. Additionally, the WDR82 subunit of the SET1A/B complexes is proposed to directly interact with the Serine 5-phosphorylated C-terminal heptad repeat domain of initiating RNAPII [95,96]. This interaction could elevate H3K4me3 at gene promoters during

transcription initiation, either by stabilising SET1A/B binding, or by stimulating their catalytic activity (Fig. 3B). In agreement with this possibility, there is a widespread requirement for WDR82 in the acquisition of H3K4me3 during lipopolysaccharide-induced gene induction in macrophages [94]. However, the removal of WDR82 does not result in a global loss of H3K4me3 from CGI-associated gene promoters in the steady state, suggesting that WDR82 is not required for maintaining H3K4me3 at CGI promoters [94].

In addition to the initiation of gene transcription, the transition from initiation to elongation has also been implicated in the deposition of H3K4me3. MLL1 directly interacts with PAF1 (Polymerase Associated Factor 1), a core component of the PAF1 complex which facilitates transcription elongation [71,97,98]. Transient recruitment of



MLL1 during the transition from initiation to elongation could lead to H3K4me3 deposition at actively transcribed genes (Fig. 3B). However, immunofluorescence experiments have shown that there is a minimal overlap in the localisation of MLL1 and PAF1, suggesting that this interaction may only occur at a small number of loci *in vivo* [71]. Additionally, it seems unlikely that MLL1/2 are responsible for elevated H3K4me3 at actively transcribed genes given that they broadly occupy CGI-associated gene promoters regardless of their transcriptional state [25,72]. While the precise mechanisms governing H3K4me3 enrichment at actively transcribed CGI-associated gene promoters remain poorly defined, it seems likely that the process of active transcription is required. Once H3K4me3 is elevated at actively transcribed gene promoters, this could initiate a simple feedback mechanism whereby the SET1A/B complexes are stabilised at transcribed genes *via* multivalent binding to non-methylated CpG DNA and H3K4me3, further amplifying H3K4me3 and SET1A/B complex occupancy (Fig. 3C).

### 3.2. Regulating the activity of the H3K4 HMTs at CpG island-associated gene promoters

Whilst CGIs play a central role in specifying how the H3K4 HMTs identify target sites in the genome, it is also clear that other features of CGI chromatin can significantly affect H3K4me3 deposition. For example, mono-ubiquitylation of histone H2B on lysine 120 (H2BK120ub1) stimulates the enzymatic activity of both MLL1 and SET1A on mononucleosomal substrates *in vitro*, and H2BK120ub1 is a prerequisite for H3K4me3 deposition on reconstituted polynucleosomal chromatin templates [99–102]. Recent cryo-EM structures of H3K4 methyltransferases bound to ubiquitylated nucleosomes have revealed possible mechanisms for this regulation. Structures of the yeast SET1C complex bound to both unmodified and H2BK120ub1 nucleosomes revealed that a conserved arginine-rich motif in the N-SET domain of SET1C binds to H2BK120ub1 and induces a conformational change which may activate HMT activity [103,104]. Although this motif is conserved in vertebrate SET1A/B, it is not retained in MLL proteins. Instead, studies of an MLL1 complex bound to unmodified and H2BK120ub1 nucleosomes suggested that activation by H2BK120ub1 may arise from restriction of MLL1 binding geometry and enhanced affinity for nucleosome substrates [100].

H3K4 HMT activity is clearly regulated by H2BK120ub1 *in vitro*, however the relevance of this for H3K4me3 deposition *in vivo* remains poorly defined. Like H3K4me3, the level of H2BK120ub1 within genes correlates with gene activity [105,106]. However, the distribution of H2BK120ub1 within genes differs significantly from that of H3K4me3. Whilst H3K4me3 is present at gene promoters and over CGI elements, H2BK120ub1 is distributed throughout the transcribed regions of genes with only modest enrichment towards the 5' end [105,106] (Fig. 3D). The mechanism by which H2BK120ub1 might contribute to the formation of H3K4me3 domains *in vivo* therefore remains unclear. *In vitro* observations suggest H2BK120ub1 may simply be a prerequisite for H3K4me3 deposition, and depletion of H2BK120ub1 has been associated with reduced H3K4me3 *in vivo* [101,102,105]. However, H3K4me3 seems to occur independently of H2BK120ub1 in some differentiated cell types [107]. It is also possible that H2BK120ub1 shapes the level or distribution of H3K4me3. Indeed, loss of H2BK120ub1 following knockout of the H2B E3 ligase component RNF40 is associated with reduced H3K4me3 predominantly on the genic side of actively transcribed gene promoters [105].

Whilst H2BK120ub1 promotes H3K4 methylation, other histone modifications can prevent H3K4me3 deposition. An inverse relationship between asymmetrically dimethylated histone H3 arginine 2 (H3R2me2a) and H3K4me3 has been observed in several instances; increased H3R2me2a following overexpression of the methyltransferase PRMT6 is associated with a decrease in bulk H3K4me3, whilst knockout of PRMT6 leads to reduced H3R2me2a and elevated H3K4me3 at gene promoters [108,109]. Interestingly, occupancy of MLL1 and the H3K4

HMT subunit WDR5 are also inversely correlated with H3R2me2a, suggesting that H3R2me2a may prevent H3K4 HMT binding [109–111]. Given the close proximity of H3K4 and H3R2 in the histone H3 tail, many H3K4me3-binding proteins make extensive contacts with H3R2 and its methylation has been shown to alter their binding affinity in a number of instances [112–117]. Notably, H3R2me2a could affect binding of the SET1A/B complexes since the PHD domain of CFP1 has a lower affinity for doubly modified H3R2me2aK4me3 peptides than for singly modified H3K4me3 peptides *in vitro* [113]. Furthermore, MLL/SET1 complexes are less active towards H3 peptides already carrying R2me2a, suggesting that H3R2me2a may prevent H3K4me3 deposition by limiting both the binding and activity of H3K4 HMTs [110,111]. The functional relevance of the crosstalk between H2R2me2a and H3K4me3 is not clear; H3R2me2a has generally been associated with transcription repression, however H3R2me2a colocalises with H3K4me3 at actively transcribed gene promoters [108,109,111]. It has been proposed in this context that H3R2me2a functions to constrain or fine-tune H3K4me3 levels and transcriptional activity, however further work is required to explore this possibility.

### 3.3. Shaping H3K4me3 after its deposition

The observed distribution of most histone modifications is dictated by mechanisms that regulate both deposition and removal. In contrast to some other trimethyl lysine modifications, which are stable over several cell divisions, H3K4me3 is relatively dynamic, with a half-life of 7 h [118]. In agreement with the dynamic nature of H3K4me3, there are four well-characterised H3K4me3 demethylases in mammals: KDM5A-D [119–123]. While these enzymes have been implicated in developmental processes and are perturbed in a range of cancers, there remains a relatively limited understanding of how they function to regulate H3K4 methylation genome-wide [124–126]. All four KDM5 family members contain a DNA-binding ARID domain and multiple PHD domains that may be responsible for chromatin binding [125,127,128]. The third PHD domain of KDM5A binds H3K4me3, which may allow for direct recruitment of H3K4 demethylase activity to substrate chromatin [129]. *In vivo*, KDM5B binds to sites in the genome that are enriched for H3K4me3, implicating it in the dynamic turnover of H3K4me3 at these loci (Fig. 3E, left) [130]. KDM5 proteins also appear to act near gene promoters to prevent H3K4me3 spreading into gene bodies, which may serve to focus H3K4me3 specifically at CGI promoters (Fig. 3E, centre) [130]. In addition, KDM5B has been shown to associate with the H3K36me3-binding protein MRG15, which may target it to intragenic regions to remove aberrantly deposited H3K4me3 (Fig. 3E, right) [131]. Interestingly, the PHD1 domains of KDM5A and KDM5B have been shown to allosterically activate these enzymes upon binding their own product, H3K4me0 [132–134]. This positive feedback mechanism may drive propagation of demethylation along nucleosomes to remove large H3K4me3 domains upon switching to a transcriptionally inactive state, or serve to maintain methylation-free chromatin domains [131,135]. Demethylases could also contribute to the enrichment of H3K4me3 at actively transcribed promoters by counteracting H3K4me3 accumulation at lowly transcribed loci. Our understanding of how demethylases shape H3K4me3 is still developing and further work is required to define their role at CGI-associated gene promoters.

## 4. How does H3K4me3 influence CpG island chromatin and gene expression?

### 4.1. CpG islands, H3K4me3, and DNA methylation

A key feature of CGI-associated gene promoters is that they tend to remain free of DNA methylation [38,41]. Given that the abnormal methylation of CGI gene promoters is associated with stable transcriptional repression, for example in cancer, it has been proposed that the

non-methylated state of CGIs contributes to the maintenance of a chromatin environment that is responsive to gene regulatory cues [136–139]. However, it is unknown how CGIs remain refractory to DNA methylation when the surrounding genome is pervasively methylated. Because DNA methylation and H3K4me3 are almost entirely mutually exclusive, it was proposed that H3K4 methylation may prevent the methylation of CGIs [52,140–143]. Indeed, a direct inverse relationship between H3K4me3 and DNA methylation has now been observed in a number of studies [48,53,144,145]. In mice, perturbation of H3K4me3 is associated with increased DNA methylation at certain genomic loci, while the expression of a mammalian DNA methyltransferase in yeast, an organism devoid of DNA methylation, results in the accumulation of aberrant DNA methylation at H3K4me3-depleted regions [18,141,144,146–148]. This atypical accumulation of DNA methylation in yeast is further enhanced in the absence of H3K4me3, suggesting that DNA methyltransferase activity is blocked at sites marked by H3K4me3 [141,148].

A family of DNA methyltransferases (DNMTs) catalyses the methylation of CG dinucleotides throughout the mammalian genome. DNMT3A/B, aided during early development by their non-catalytic cofactor DNMT3L, primarily catalyse *de novo* DNA methylation. DNMT1 is a maintenance methyltransferase which ensures that DNA methylation patterns are re-established after cell division (reviewed in Gowher and Jeltsch [149]). Detailed structural and biochemical studies of the DNMTs have uncovered potential mechanisms by which H3K4me3 could counteract the deposition of DNA methylation at CGIs. The ADD domains of DNMT3A/B and DNMT3L preferentially bind to the N-terminal tail of histone H3 when H3K4 is non-methylated [150–153]. Rendering the DNMT3A ADD domain insensitive to the H3K4 methylation state in mouse ESCs results in increased DNMT3A binding and CGI-associated DNA methylation at sites with H3K4me3 [153]. This suggests that H3K4 methylation precludes association of the *de novo* DNA methyltransferases with CGIs and therefore prevents the acquisition of DNA methylation. However, the relevance of the ADD domains for correct genomic targeting *in vivo* has been debated [154,155].

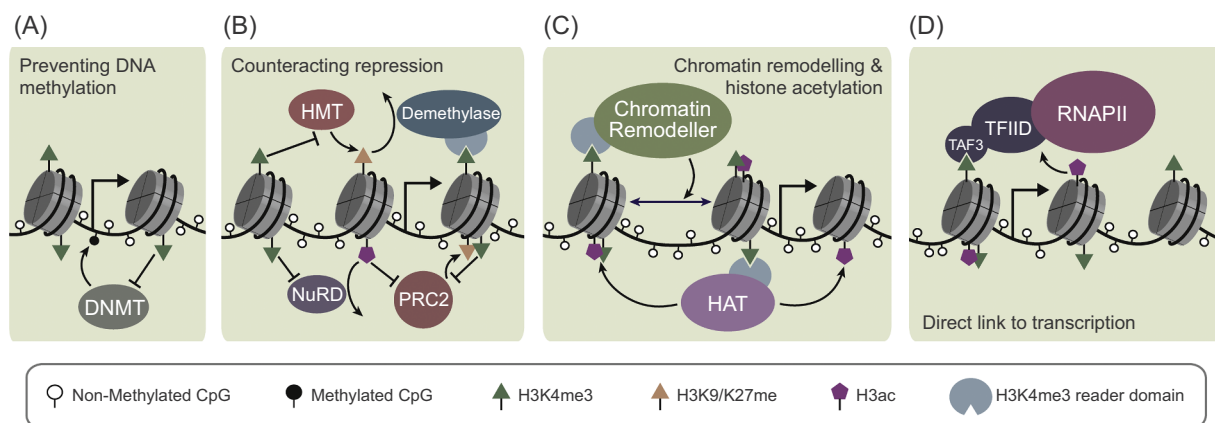
It has also been proposed that binding to H3 regulates the enzymatic activity of DNMT3A, which is thought to reside in an autoinhibited state when not bound to chromatin. Release of this autoinhibition is

dependent on DNMT3A association with H3K4me0 [152,156,157]. This could provide an additional layer of protection against DNA methylation at CGIs, since aberrant binding of DNMT3A to H3K4me3-containing chromatin would fail to release its autoinhibition. By preventing DNA methylation at CGIs (Fig. 4A), the presence of H3K4me3 could ensure that they are maintained in a transcriptionally permissive state.

#### 4.2. Does H3K4me3 counteract transcriptional repression?

ZF-CxxC domain-containing proteins are present in protein complexes associated with both transcriptional activation and repression. It is therefore essential that the CGI-associated gene promoters of actively transcribed genes are resistant to the activities of repressive complexes that can bind to these loci. In agreement with this, there is evidence that H3K4me3 may counteract chromatin states associated with transcriptional repression. For example, H3K4me3 has been implicated in inhibiting the deposition of repressive, heterochromatin-associated histone modifications, most notably methylation of H3K9 and H3K27. H3K4me3 has been proposed to regulate H3K9 methylation by inhibiting binding of the H3K9 HMTs, preventing its deposition, and also by stabilising binding of H3K9 demethylases, thereby promoting its removal from gene promoters (Fig. 4B) [158–173].

The relationship between H3K4me3 and H3K27me3, a histone modification catalysed by the Polycomb repressive system, has also been extensively studied (Reviewed in Voigt et al. [174]). From *in vitro* experiments, it is clear that H3K4me3 inhibits H3K27me3 deposition by Polycomb-repressive complex 2 (PRC2) on the same N-terminal tail of histone H3. However, the mechanism for this inhibition remains unclear [175–178]. There is also *in vitro* evidence that binding of the NuRD (Nucleosome Remodelling and Deacetylase) complex to histone H3 peptides is reduced in the presence of H3K4me3 [86,172,179–184]. Since NuRD-mediated deacetylation promotes PRC2 engagement with chromatin, this may be a mechanism whereby H3K4me3 indirectly counteracts the formation of transcriptionally repressive Polycomb domains (Fig. 4B) [185]. Somewhat counterintuitively, however, ChIP-seq in mammalian cells has shown that the localisation of H3K27me3 and H3K4me3 can extensively overlap at CGIs, particularly in ESCs. This so-called ‘bivalent’ chromatin has been proposed to retain genes in a repressed, but ‘poised’, configuration to enable rapid



**Fig. 4.** The relationship between H3K4me3 and transcriptionally permissive chromatin.

(A) H3K4me3 has been proposed to counteract methylation of CpG islands by preventing both binding and activation of the *de novo* DNA methyltransferases (DNMTs). The position of the transcription start site is illustrated as an arrow in each panel and a legend for the chromatin modifications is shown below the panels. (B) H3K4me3 counteracts acquisition of repressive chromatin modifications by preventing binding of histone methyltransferases (HMTs) while promoting removal of repressive methylation by stabilising binding of demethylases. In addition to inhibiting its catalytic activity, H3K4me3 may also inhibit PRC2 by preventing binding of NuRD and hence indirectly reduce PRC2 association with chromatin.

(C) Histone acetyltransferases (HATs) and chromatin remodellers bind to H3K4me3 and contribute to an accessible and transcriptionally permissive chromatin architecture at CGI-associated gene promoters.

(D) The TAF3 subunit of TFIID binds H3K4me3 and promotes pre-initiation complex formation, providing a direct link to transcription. This interaction is enhanced by histone acetylation.

activation during differentiation [186–188]. There is, however, currently little evidence for the role of bivalent chromatin in supporting gene activation [24,25,72,189].

#### 4.3. Regulation of histone acetylation and chromatin remodelling by H3K4me3

Biochemical approaches aimed at identifying factors that interact with the N-terminal tail of histone H3 have identified a number of proteins that specifically bind to, or are excluded by, H3K4me3 [30,86,89,190]. These include several protein complexes associated with histone acetylation and chromatin remodelling.

Much like H3K4me3, histone acetylation is found at active gene promoters [9,191]. However, in contrast to histone methylation, which does not affect the charge of lysine residues, histone acetylation neutralises their positive charge, weakening histone tail-DNA interactions and counteracting chromatin compaction [192]. It is therefore likely that the presence of histone acetylation at CGI-associated gene promoters increases their accessibility and therefore contributes to a transcriptionally permissive chromatin state. The close relationship between H3K4me3 and histone acetylation is exemplified by the SAGA (Spt-Ada-Gcn5 Acetyltransferase) chromatin-modifying complex, which has been implicated in transcriptional activation [193]. The SGF29 (SAGA-associated Factor 29) subunit of the SAGA complex binds to H3K4me3 via its tandem Tudor domains [89,194]. This is important for the association of the SAGA complex with chromatin and acetylation of its target loci. However, despite there being a major reduction in histone acetylation upon H3K4me3 perturbation *in vivo*, there is evidence that the relationship between H3K4me3 and histone acetylation may not be so straightforward [195,196]. The mammalian ING proteins (INhibitor of Growth 1–5) each contain a PHD domain that preferentially binds to H3K4me3 [197–204]. While ING3–5 associate with histone acetyltransferase complexes, ING1 and ING2 associate with the SIN3 histone deacetylase complex, which removes histone acetylation and can therefore counteract gene expression [205–209]. It is possible that, by recruiting both histone acetyltransferases and histone deacetylases to CGI-associated gene promoters, H3K4me3 contributes to the dynamics and/or distribution of histone acetylation [210–212]. Further work is required to better define how histone acetylation is regulated at CGI gene promoters, although H3K4me3 appears to be central to this process.

Nucleosome occupancy and phasing at gene promoters have been widely implicated in regulating gene expression (reviewed in Lai and Pugh and Brahma and Henikoff [213,214]). Actively transcribed CGI-associated gene promoters, which have high H3K4me3, are characterised by dynamic nucleosomes in which the underlying DNA is more accessible compared to the CGI-associated promoters of lowly transcribed genes [215,216]. It has been proposed that this chromatin state relies on H3K4me3 and its capacity to stabilise the binding of chromatin remodelling enzymes [217,218]. For example, BPTF (Bromodomain and PHD finger-containing Transcription Factor), a subunit of the NURF (NUcleosome Remodelling Factor) chromatin-remodelling complex, has a PHD domain that binds to H3K4me3 [219,220]. NURF utilises the energy derived from ATP-hydrolysis to slide nucleosomes on DNA, and has therefore been implicated in transcriptional regulation [221]. BPTF association with H3K4me3 is required for appropriate NURF complex binding and remodelling activity on chromatin *in vivo* [217,219]. Interestingly, BPTF also has a bromodomain that binds acetylated histone H4 tails and enables it to engage in multivalent interactions with chromatin carrying both H3K4me3 and H4 acetylation [222]. These modifications at actively transcribed genes may therefore reinforce a dynamic and accessible chromatin state to support ongoing gene transcription. In addition to BPTF/NURF, the ATP-dependent chromatin remodelling enzyme CHD1 (Chromodomain Helicase DNA binding protein 1) also binds to H3K4me3 via its double chromodomains [223,224]. H3K4me3 has been proposed to enhance the

nucleosome remodelling activity of CHD1, which is thought to contribute to the maintenance of a transcriptionally-permissive chromatin state at gene promoters [225,226]. Together these observations suggest that H3K4me3 and histone acetylation may promote gene expression, at least in part, by recruiting chromatin remodelling enzymes to actively transcribed CGI-associated gene promoters (Fig. 4C).

#### 4.4. H3K4me3 and a direct link to the transcriptional machinery

Most of the characterised H3K4me3 reader proteins associate with activities that modify or remodel chromatin, and many of these have proposed roles in gene regulation. However, there is also evidence that H3K4me3 may directly affect the transcriptional machinery itself. The TAF3 subunit of the TFIID general transcription factor complex contains a PHD domain that associates with H3K4me3 and promotes pre-initiation complex (PIC) formation on *in vitro* transcription templates [86,114,115]. In a recent model of RNAPII engagement with chromatin, it was proposed that the association of TFIID with nucleosomes precedes engagement of TBP (TATA-Binding Protein) with DNA [227]. Therefore, additional mechanisms that support nucleosome binding by TFIID, like recognition of H3K4me3, could be more important for initiation of RNAPII-mediated gene transcription than previously appreciated [228]. Indeed, TAF3 binding to chromatin *in vivo* significantly overlaps with H3K4me3 at CGI-associated gene promoters, and binding to H3K4me3 is essential for recruitment of TBP and the appropriate induction of many TAF3 target genes [229,230]. It is therefore possible that, in addition to contributing to a transcriptionally permissive chromatin environment, H3K4me3 may directly affect RNAPII function by mediating PIC formation at CGI-associated gene promoters (Fig. 4D). Interestingly, binding of TFIID to H3K4me3-containing chromatin is also enhanced by histone acetylation, another key feature of actively transcribed CGI-associated gene promoters [114,230] (Section 4.3.).

#### 4.5. Does H3K4me3 regulate gene expression?

As we have previously described, H3K4me3 correlates with gene transcription, contributes to a chromatin environment that is seemingly permissive to gene transcription, and has a direct link to the transcriptional machinery (Figs. 2, 4). It has therefore been proposed that H3K4me3 directly contributes to gene transcription [114,229]. In support of this, it has been shown in some contexts that efficient transcription *in vitro* is dependent on H3K4 methylation, though in others its contribution is minimal [101,225,229,231,232]. However, there remains little evidence directly linking the loss of H3K4me3 to major genome-wide alterations in gene expression *in vivo*. For example, yeast strains completely lacking SET1C and H3K4me3 are viable and have only mild transcriptional defects. This may be because yeast TAF3 lacks the H3K4me3-binding PHD domain which has been proposed to stimulate PIC formation in mammalian cells [13,233–243]. Studying the requirement of H3K4me3 for gene expression in vertebrates is complicated by the fact that there are multiple H3K4me3 HMTs, which may have overlapping function and also methyltransferase-independent functions [83,244]. While depletion of components of the SET1A/B complexes in vertebrate cells results in clear defects in gene expression, these changes are subtle and often manifest as decreases in gene expression at some genes and increases in gene expression at others, and there is limited correlation between gene expression changes and alterations in H3K4me3 [61,94,195,245,246]. This suggests that, at least in the steady state, alterations in H3K4me3 do not have a predictable effect on gene expression [247].

It has been proposed that the acquisition of H3K4me3 may be important for gene induction, for example during differentiation [83,248,249]. However, perturbation of the SET1A/B complexes during LPS-mediated gene induction in macrophages and DNA damage-induced gene expression in ESCs does not result in widespread transcriptional defects, despite reduced H3K4me3 levels at activated genes

[94,195]. Similarly, the removal of MLL2 in ESCs reduces H3K4me3 at lowly transcribed or inactive CGI-associated genes but does not affect the induction of the majority of these genes during retinoic acid-induced differentiation, or differentiation into embryoid bodies [24,25,72,189]. Therefore, at least under the kinetic parameters examined thus far, there remains little evidence that H3K4me3 acquisition is essential for gene induction.

Single-cell measurements have shown that the transcription of genes associated with broad domains of H3K4me3 is more consistent, or less varied, between cells [250]. This transcriptional consistency appears to be dependent on the presence of H3K4me3, suggesting that H3K4me3 somehow promotes more consistent transcription. One could imagine that cell-to-cell variation in gene expression could be deleterious to cell fate decisions during differentiation and to coordinated cellular functions in tissues that require homogeneous gene expression. This may explain why SET1A/B and MLL1/2 complexes are essential for normal embryonic development, but their removal in cell culture systems results in seemingly modest effects on gene expression [26,251–254]. It will be important in future work to determine whether the developmental defects that arise following removal of SET1A/B and MLL1/2 are due to alterations in H3K4me3 and transcriptional consistency.

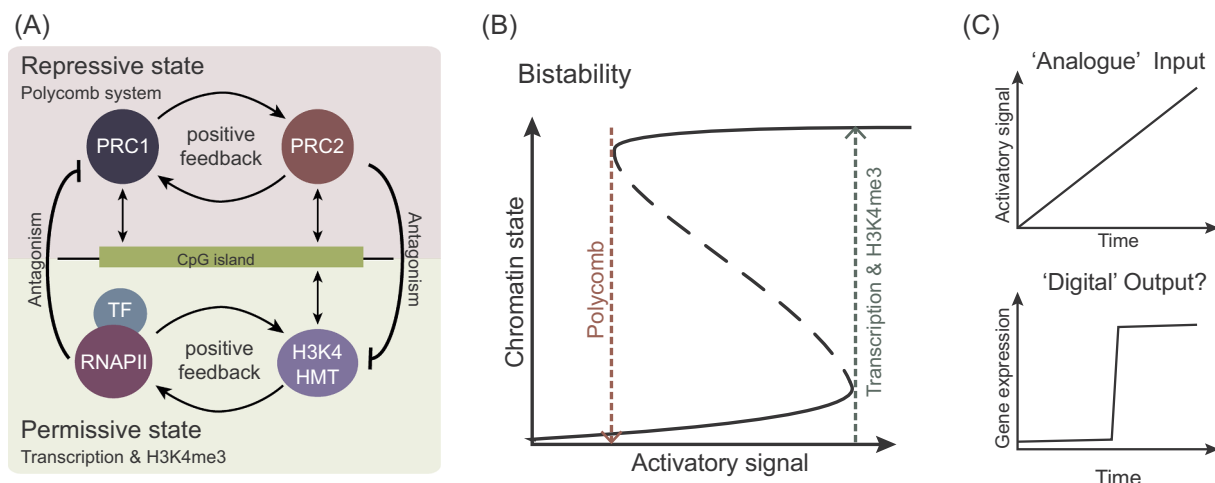
#### 4.6. Does H3K4me3 contribute to chromatin bistability at CGI-associated promoters?

Although the precise role of H3K4me3 in vertebrate gene expression has yet to be fully defined, its proposed contribution to transcriptional consistency suggests that this system may regulate gene expression in a manner that is masked by ensemble measurements and may only be evident when assayed with appropriate kinetic measurements or in single-cell experiments. We and others have previously proposed that CGIs may create bistable chromatin states at gene promoters to regulate gene expression [138,174,255–263]. In the context of such a system, we propose that repressive Polycomb chromatin states maintained by

feedback mechanisms at inactive genes constrain gene activation signals until an appropriate gene induction threshold is reached. In contrast, following productive initiation of transcription, the establishment of permissive chromatin states potentiates transcription, provides memory that the gene was recently transcribed, and allows for gene expression to be maintained even when levels of activatory signal drop below the threshold required for initial activation (Fig. 5A, B). We envisage that this switching between chromatin states could allow analogue gene induction signals to be converted into digital gene expression outputs through the behaviour of the underlying chromatin at CGIs (Fig. 5C).

In the context of this model for chromatin bistability at CGIs, we envisage that the H3K4 HMT complexes dynamically engage with, or sample, CGIs via their ZF-CxxC domains, but this is insufficient for stable occupancy and high level H3K4me3 deposition. During gene activation, transcription factor-dependent or cotranscriptional recruitment of H3K4 HMTs would lead to increased H3K4me3 at target gene CGIs. This elevated H3K4me3 associated with gene activation could be recognised by PHD domains in H3K4 HMT complexes to support multivalent binding that would further elevate H3K4me3 at transcribed CGI-associated gene promoters (Fig. 3). Elevated H3K4me3 would in turn promote gene transcription, creating a feedback loop that could also counteract repressive chromatin states (Figs. 4, 5A). Given that vertebrate gene expression is inherently stochastic, a permissive H3K4me3 chromatin state at transcribed gene promoters may support uniformity in gene expression and provide chromatin-based memory of recent transcriptional events [264].

Recent mathematical models integrating the known activities of complexes that drive permissive and repressive chromatin states at CGIs have provided some evidence that bistable chromatin states are likely to exist, and that such systems could potentially modulate transcriptional behaviours in a manner that is consistent with the ideas described here and previously [138,174,255–263,265]. In the context of these emerging models, it is important to highlight that previous work



**Fig. 5.** Do CpG island-associated gene promoters and chromatin bistability shape transitions in gene expression?

(A) Both the Polycomb system (PRC1 and PRC2) and H3K4 histone methyltransferases (HMTs) can engage with or 'sample' CGIs via ZF-CxxC domain-containing proteins. Self-reinforcing positive feedback loops and mutual antagonism inherent to these systems could create bistable chromatin states at CGIs.

(B) A schematic illustrating the proposed bistable chromatin state at CpG island-associated gene promoters. The transcriptionally repressive Polycomb state is sustained by feedback mechanisms that also antagonise the permissive state (A). We envisage that this Polycomb dependent repressive state constrains transcription until gene activation signals reach a threshold where inhibition is overcome and transcription initiates effectively. This would then result in a transcription-dependent switch to a permissive chromatin state, characterised by H3K4me3 and transcriptional activity (green arrow). This transcriptionally permissive state is maintained through self-reinforcing feedback mechanisms that also antagonise the Polycomb state (A). The permissive state is thereby maintained unless the activatory signal drops below a certain level where feedback can no longer be sustained and transcription does not persist. This would then cause a switch back to the repressive Polycomb chromatin state (red arrow).

(C) The interaction between regulatory signals and bistable chromatin states (A,B) at CGI-associated gene promoters could shape gene expression transitions such that graded analogue gene regulatory inputs are translated into switch-like digital gene expression outputs.



examining the effects of H3K4me3 perturbation on gene induction relied on ensemble transcriptional measures. Such measures are incapable of distinguishing between analogue and digital transcriptional behaviours and lack the time resolution required to capture kinetic contributions to these gene expression transitions. Therefore, future studies will require suitably quantitative and time-resolved methods to examine the contribution of CGIs and H3K4me3 to gene expression and to test whether chromatin bistability may contribute to appropriate gene expression transitions.

## 5. Summary

It is clear that H3K4 methylation is intrinsically linked to CGIs in vertebrate chromatin. H3K4me3 deposition at CGIs appears to shape a chromatin environment around gene promoters that supports communication with the transcriptional machinery. However, the effects of perturbing the H3K4me3 system on gene expression are not straightforward and the mechanisms governing the role of CGI-associated H3K4me3 in the regulation of gene expression remain poorly understood. Due to the potentially bistable nature of CGI chromatin, we envisage that single cell technologies that capture the dynamic behaviours of gene expression will be required to enable further investigation into how H3K4me3 contributes to gene regulation.

## CrediT authorship contribution statement

**Amy L. Hughes:** Conceptualization, Writing - original draft, Writing - review & editing. **Jessica R. Kelley:** Conceptualization, Writing - original draft, Writing - review & editing. **Robert J. Klose:** Conceptualization, Writing - original draft, Writing - review & editing, Supervision.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

We would like to thank Neil Blackledge, Emilia Dimitrova and Emma Smith for critical reading of the manuscript.

## Funding

Work in the Klose lab is supported by the Wellcome Trust (209400/Z/17/Z), the European Research Council (681440), and the Lister Institute of Preventive Medicine. ALH is supported by a Wellcome Trust studentship (203829/Z/16/A) and JRK is supported by the Oxford-Wolfson Marriot Graduate Scholarship.

## References

- [1] Y. Lorch, J.W. LaPointe, R.D. Kornberg, Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones, *Cell* 49 (1987) 203–210, [https://doi.org/10.1016/0092-8674\(87\)90561-7](https://doi.org/10.1016/0092-8674(87)90561-7).
- [2] B. Li, M. Carey, J.L. Workman, The role of chromatin during transcription, *Cell* 128 (2007) 707–719, <https://doi.org/10.1016/j.cell.2007.01.015>.
- [3] M. Levine, C. Cattoglio, R. Tjian, Looping back to leap forward: transcription enters a new era, *Cell* 157 (2014) 13–25, <https://doi.org/10.1016/j.cell.2014.02.009>.
- [4] S. Venkatesh, J.L. Workman, Histone exchange, chromatin structure and the regulation of transcription, *Nat. Rev. Mol. Cell Biol.* 16 (2015) 178–189, <https://doi.org/10.1038/nrm3941>.
- [5] S.B. Rothbart, B.D. Strahl, Interpreting the language of histone and DNA modifications, *Biochim. Biophys. Acta - Gene Regul. Mech.* 1839 (2014) 627–643, <https://doi.org/10.1016/j.bbagen.2014.03.001>.
- [6] T. Zhang, S. Cooper, N. Brockdorff, The interplay of histone modifications – writers that read, *EMBO Rep.* 16 (2015) 1467–1481, <https://doi.org/10.15252/embr.201504945>.
- [7] L.M. Soares, P.C. He, Y. Chun, H. Suh, T.S. Kim, S. Buratowski, Determinants of histone H3K4 methylation patterns, *Mol. Cell* 68 (2017) 773–785.e6, <https://doi.org/10.1016/j.molcel.2017.10.013>.
- [8] R. Schneider, A.J. Bannister, F.A. Myers, A.W. Thorne, C. Crane-Robinson, T. Kouzarides, Histone H3 lysine 4 methylation patterns in higher eukaryotic genes, *Nat. Cell Biol.* 6 (2004) 73–77, <https://doi.org/10.1038/ncb1076>.
- [9] B.E. Bernstein, M. Kamal, K. Lindblad-Toh, S. Bekiranov, D.K. Bailey, D.J. Huebert, S. McMahon, E.K. Karlsson, E.J. Kulbokas, T.R. Gingeras, S.L. Schreiber, E.S. Lander, Genomic maps and comparative analysis of histone modifications in human and mouse, *Cell* 120 (2005) 169–181, <https://doi.org/10.1016/j.cell.2005.01.001>.
- [10] N.D. Heintzman, R.K. Stuart, G. Hon, Y. Fu, C.W. Ching, R.D. Hawkins, L.O. Barrera, S. Van Calcar, C. Qu, K.A. Ching, W. Wang, Z. Weng, R.D. Green, G.E. Crawford, B. Ren, Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome, *Nat. Genet.* 39 (2007) 311–318, <https://doi.org/10.1038/ng1966>.
- [11] N.D. Heintzman, G.C. Hon, R.D. Hawkins, P. Kheradpour, A. Stark, L.F. Harp, Z. Ye, L.K. Lee, R.K. Stuart, C.W. Ching, K.A. Ching, J.E. Antosiewicz-Bourget, H. Liu, X. Zhang, R.D. Green, V.V. Lobanenkov, R. Stewart, J.A. Thomson, G.E. Crawford, M. Kellis, B. Ren, Histone modifications at human enhancers reflect global cell-type-specific gene expression, *Nature* 459 (2009) 108–112, <https://doi.org/10.1038/nature07829>.
- [12] S.D. Briggs, M. Bryk, B.D. Strahl, W.L. Cheung, J.K. Davie, S.Y.R. Dent, F. Winston, C. David Allis, Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in *Saccharomyces cerevisiae*, *Genes Dev.* 15 (2001) 3286–3295, <https://doi.org/10.1101/gad.940201>.
- [13] H. Santos-Rosa, R. Schneider, A.J. Bannister, J. Sherriff, B.E. Bernstein, N.C.T. Emre, S.L. Schreiber, J. Mellor, T. Kouzarides, Active genes are tri-methylated at K4 of histone H3, *Nature* 419 (2002) 407–411, <https://doi.org/10.1038/nature01080>.
- [14] A. Shilatifard, The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis, *Annu. Rev. Biochem.* 81 (2012) 65–95, <https://doi.org/10.1146/annurev-biochem-051710-134100>.
- [15] J.H. Lee, D.G. Skalnik, CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex, *J. Biol. Chem.* 280 (2005) 41725–41731, <https://doi.org/10.1074/jbc.M508312200>.
- [16] J.-H. Lee, C.M. Tate, J.-S. You, D.G. Skalnik, Identification and characterization of the human Set1B histone H3-Lys 4 methyltransferase complex, *J. Biol. Chem.* 282 (2007) 13419–13428, <https://doi.org/10.1074/jbc.M609809200>.
- [17] K.T. Fitzgerald, M.O. Diaz, MLL2: A new mammalian member of the trx/MLL family of genes, *Genomics* 59 (1999) 187–192, <https://doi.org/10.1006/geno.1999.5860>.
- [18] T.A. Milne, S.D. Briggs, H.W. Brock, M.E. Martin, D. Gibbs, C.D. Allis, J.L. Hess, MLL targets SET domain methyltransferase activity to Hox gene promoters, *Mol. Cell* 10 (2002) 1107–1117, [https://doi.org/10.1016/S1097-2765\(02\)00741-4](https://doi.org/10.1016/S1097-2765(02)00741-4).
- [19] T. Nakamura, T. Mori, S. Tada, W. Krajewski, T. Rozovskaia, R. Wassell, G. Dubois, A. Mazo, C.M. Croce, E. Canaani, ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation, *Mol. Cell* 10 (2002) 1119–1128, [https://doi.org/10.1016/S1097-2765\(02\)00740-2](https://doi.org/10.1016/S1097-2765(02)00740-2).
- [20] Y.W. Cho, T. Hong, S.H. Hong, H. Guo, H. Yu, D. Kim, T. Guszczynski, G.R. Dressler, T.D. Copeland, M. Kalkum, K. Ge, PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex, *J. Biol. Chem.* 282 (2007) 20395–20406, <https://doi.org/10.1074/jbc.M701574200>.
- [21] J.E. Lee, C. Wang, S. Xu, Y.W. Cho, L. Wang, X. Feng, A. Baldrige, V. Sartorelli, L. Zhuang, W. Peng, K. Ge, H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation, *Elife* (2013) 2013, <https://doi.org/10.7554/eLife.01503>.
- [22] D. Hu, X. Gao, M.A. Morgan, H.-M. Herz, E.R. Smith, A. Shilatifard, The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 mono-methylases at enhancers, *Mol. Cell Biol.* 33 (2013) 4745–4754, <https://doi.org/10.1128/mcb.01181-13>.
- [23] S. Glaser, S. Lubitz, K.L. Loveland, K. Ohbo, L. Robb, F. Schwenk, J. Seibler, D. Roellig, A. Kranz, K. Anastasiadis, A.F. Stewart, The histone 3 lysine 4 methyltransferase, MLL2, is only required briefly in development and spermatogenesis, *Epigenetics and Chromatin* 2 (2009) 5, <https://doi.org/10.1186/1756-8935-2-5>.
- [24] D. Hu, A.S. Garruss, X. Gao, M.A. Morgan, M. Cook, E.R. Smith, A. Shilatifard, The MLL2 branch of the COMPASS family regulates bivalent promoters in mouse embryonic stem cells, *Nat. Struct. Mol. Biol.* 20 (2013) 1093–1097, <https://doi.org/10.1038/nsmb.2653>.
- [25] S. Denisov, H. Hofmeister, H. Marks, A. Kranz, G. Ciotta, S. Singh, K. Anastasiadis, H.G. Stunnenberg, A.F. Stewart, MLL2 is required for H3K4 tri-methylation on bivalent promoters in embryonic stem cells, whereas MLL1 is redundant, *Dev* 141 (2014) 526–537, <https://doi.org/10.1242/dev.102681>.
- [26] A.S. Bledau, K. Schmidt, K. Neumann, U. Hill, G. Ciotta, A. Gupta, D.C. Torres, J. Fu, A. Kranz, A.F. Stewart, K. Anastasiadis, The H3K4 methyltransferase Setd1a is first required at the epiblast stage, whereas Setd1b becomes essential after gastrulation, *Dev* 141 (2014) 1022–1035, <https://doi.org/10.1242/dev.098152>.
- [27] L. Fang, J. Zhang, H. Zhang, X. Yang, X. Jin, L. Zhang, D.G. Skalnik, Y. Jin, Y. Zhang, X. Huang, J. Li, J. Wong, H3K4 methyltransferase set1a is a key Oct4 coactivator essential for generation of Oct4 positive inner cell mass, *Stem Cells* 34 (2016) 565–580, <https://doi.org/10.1002/stem.2250>.
- [28] R. van Nuland, A.H. Smits, P. Pallaki, P.W.T.C. Jansen, M. Vermeulen, H.T.M. Timmers, Quantitative dissection and stoichiometry determination of the

- human SET1/MLL histone methyltransferase complexes, *Mol. Cell. Biol.* 33 (2013) 2067–2077, <https://doi.org/10.1128/mcb.01742-12>.
- [29] T. Kusch, Histone H3 lysine 4 methylation revisited, *Transcription* 3 (2012) 310–314, <https://doi.org/10.4161/trns.21911>.
- [30] K. Hyun, J. Jeon, K. Park, J. Kim, Writing, erasing and reading histone lysine methylations, *Exp. Mol. Med.* 49 (2017) e324, <https://doi.org/10.1038/emmm.2017.11>.
- [31] H.H. Ng, F. Robert, R.A. Young, K. Struhl, Targeted recruitment of Set1 histone methylase by elongating pol II provides a localized mark and memory of recent transcriptional activity, *Mol. Cell* 11 (2003) 709–719, [https://doi.org/10.1016/S1097-2765\(03\)00092-3](https://doi.org/10.1016/S1097-2765(03)00092-3).
- [32] N.J. Krogan, J. Dover, A. Wood, J. Schneider, J. Heidt, M.A. Boateng, K. Dean, O.W. Ryan, A. Golshani, M. Johnston, J.F. Greenblatt, A. Shilatifard, The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation, *Mol. Cell* 11 (2003) 721–729, [https://doi.org/10.1016/S1097-2765\(03\)00091-1](https://doi.org/10.1016/S1097-2765(03)00091-1).
- [33] A. Bochyńska, J. Lüscher-Firzlaff, B. Lüscher, Modes of interaction of KMT2 histone H3 lysine 4 methyltransferase/COMPASS complexes with chromatin, *Cells* 7 (2018) 17, <https://doi.org/10.3390/cells7030017>.
- [34] N.T. Crump, T.A. Milne, Why are so many MLL lysine methyltransferases required for normal mammalian development? *Cell. Mol. Life Sci.* 76 (2019) 2885–2898, <https://doi.org/10.1007/s00018-019-03143-z>.
- [35] C.C. Sze, A. Shilatifard, MLL3/MLL4/COMPASS family on epigenetic regulation of enhancer function and cancer, *Cold Spring Harb. Perspect. Med.* 6 (2016) a026427, <https://doi.org/10.1101/cshperspect.a026427>.
- [36] A. Jeltsch, J. Broche, P. Bashtyrykov, Molecular processes connecting DNA methylation patterns with DNA methyltransferases and histone modifications in mammalian genomes, *Genes (Basel)* 9 (2018) 566, <https://doi.org/10.3390/genes9110566>.
- [37] V.R.B. Liyanage, J.S. Jarmasz, N. Murugesan, M.R.D. Bigio, M. Rastegar, J.R. Davie, DNA modifications: function and applications in normal and disease states, *Biology (Basel)* 3 (2014) 670–723, <https://doi.org/10.3390/biology3040670>.
- [38] A. Bird, M. Taggart, M. Frommer, O.J. Miller, D. Macleod, A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA, *Cell* 40 (1985) 91–99, [https://doi.org/10.1016/0092-8674\(85\)90312-5](https://doi.org/10.1016/0092-8674(85)90312-5).
- [39] R. Lister, M. Pelizzola, R.H. Dowen, R.D. Hawkins, G. Hon, J. Tonti-Filippini, J.R. Nery, L. Lee, Z. Ye, Q.-M. Ngo, L. Edsall, J. Antosiewicz-Bourget, R. Stewart, V. Ruotti, A. H. Millar, J. A. Thomson, B. Ren, J.R. Ecker, Human DNA methylomes at base resolution show widespread epigenomic differences, *Nature* 462 (2009) 315–322, <https://doi.org/10.1038/nature08514>.
- [40] Z.D. Smith, A. Meissner, DNA methylation: roles in mammalian development, *Nat. Rev. Genet.* 14 (2013) 204–220, <https://doi.org/10.1038/nrg3354>.
- [41] M.V.C. Greenberg, D. Bourc'his, The diverse roles of DNA methylation in mammalian development and disease, *Nat. Rev. Mol. Cell Biol.* 20 (2019) 590–607, <https://doi.org/10.1038/s41580-019-0159-6>.
- [42] J. cheng Shen, W.M. Rideout, P.A. Jones, The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA, *Nucleic Acids Res.* 22 (1994) 972–976, <https://doi.org/10.1093/nar/22.6.972>.
- [43] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. Fitzhugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczyk, R. Levine, P. McEwan, K. McKernan, J. Meldrim, J.P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J.C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Showkneen, S. Sims, R.H. Waterston, R.K. Wilson, L.W. Hillier, J.D. McPherson, M.A. Marra, E.R. Mardis, L.A. Fulton, A.T. Chinwalla, K.H. Pepin, W.R. Gish, S.L. Chissoe, M.C. Wendt, K.D. Delehaunty, T.L. Miner, A. Delehaunty, J.B. Kramar, L.L. Cook, R.S. Fulton, D.L. Johnson, P.J. Minx, S.W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R.A. Gibbs, D.M. Muzny, S.E. Scherer, J.B. Bouck, E.J. Sodergren, K.C. Worley, C.M. Rives, J.H. Gorrell, M.L. Metzker, S.L. Naylor, R.S. Kucherlapati, D.L. Nelson, G.M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, D.R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, M.L. Hong, J. Dubois, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R.W. Davis, N.A. Federspiel, A.P. Abola, M.J. Proctor, B.A. Roe, F. Chen, H. Pan, J. Ramsier, H. Lehrach, R. Reinhardt, W.R. McCombie, M. De La Bastide, N. Dedhia, H. Blöcker, K. Hornischer, G. Nordsiek, R. Agarwal, L. Aravind, J.A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D.G. Brown, C.B. Burge, L. Cerutti, H.C. Chen, D. Church, M. Clamp, R.B. Copley, T. Doerks, S.R. Eddy, E.E. Eichler, T.S. Furey, J. Galagan, J.G.R. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L.S. Johnson, T.A. Jones, S. Kasif, A. Kasprzyk, S. Kennedy, W.J. Kent, P. Kitts, E.V. Koonin, I. Korf, D. Kulp, D. Lancet, T.M. Lowe, A. McLysaght, T. Mikkelsen, J.V. Moran, N. Mulder, V.J. Pollara, C.P. Ponting, G. Schuler, J. Schultz, G. Slater, A.F.A. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y.I. Wolf, K.H. Wolfe, S.P. Yang, R.F. Yeh, F. Collins, M.S. Guyer, J. Peterson, A. Felsenfeld, K.A. Wetterstrand, R.M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D.R. Cox, M.V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G.A. Evans, M. Athanasiou, R. Schultz, A. Patrinos, M.J. Morgan, Initial sequencing and analysis of the human genome, *Nature* 409 (2001) 860–921, <https://doi.org/10.1038/35057062>.
- [44] M. Gardiner-Garden, M. Frommer, CpG Islands in vertebrate genomes, *J. Mol. Biol.* 196 (1987) 261–282, [https://doi.org/10.1016/0022-2836\(87\)90689-9](https://doi.org/10.1016/0022-2836(87)90689-9).
- [45] S. Branciamore, Z.-X. Chen, A.D. Riggs, S.N. Rodin, CpG island clusters and proepigenetic selection for CpGs in protein-coding exons of HOX and other transcription factors, *Proc. Natl. Acad. Sci.* 107 (2010) 15485–15490, <https://doi.org/10.1073/pnas.1010506107>.
- [46] S. Saxonov, P. Berg, D.L. Brutlag, A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters, *Proc. Natl. Acad. Sci. U. S. A.* 103 (2006) 1412–1417, <https://doi.org/10.1073/pnas.0510310103>.
- [47] H.K. Long, D. Sims, A. Heger, N.P. Blackledge, C. Kutter, M.L. Wright, F. Grützner, D.T. Odom, R. Patient, C.P. Ponting, R.J. Klose, Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates, *Elife* (20) (2013) 13 1–19, <https://doi.org/10.7554/eLife.00348>.
- [48] R. Greenfield, A. Tabib, I. Keshet, J. Moss, O. Sabag, A. Goren, H. Cedar, Role of transcription complexes in the formation of the basal methylation pattern in early development, *Proc. Natl. Acad. Sci. U. S. A.* 115 (2018) 10387–10391, <https://doi.org/10.1073/pnas.1804755115>.
- [49] J.P. Thomson, P.J. Skene, J. Selfridge, T. Clouaire, J. Guy, S. Webb, A.R.W. Kerr, A. Deaton, R. Andrews, K.D. James, D.J. Turner, R. Illingworth, A. Bird, CpG islands influence chromatin structure via the CpG-binding protein Cfp1, *Nature* 464 (2010) 1082–1086, <https://doi.org/10.1038/nature08924>.
- [50] E. Wachter, T. Quante, C. Merusi, A. Arczewska, F. Stewart, S. Webb, A. Bird, Synthetic CpG islands reveal DNA sequence determinants of chromatin structure, *Elife* 3 (2014) e03397, <https://doi.org/10.7554/eLife.03397>.
- [51] E.M. Mendenhall, R.P. Koche, T. Truong, V.W. Zhou, B. Issac, A.S. Chi, M. Ku, B.E. Bernstein, GC-rich sequence elements recruit PRC2 in mammalian ES cells, *PLoS Genet.* 6 (2010) e1001244, <https://doi.org/10.1371/journal.pgen.1001244>.
- [52] S.A. Smallwood, S.I. Tomizawa, F. Krueger, N. Ruf, N. Carli, A. Segonds-Pichon, S. Sato, K. Hata, S.R. Andrews, G. Kelsey, Dynamic CpG island methylation landscape in oocytes and preimplantation embryos, *Nat. Genet.* 43 (2011) 811–814, <https://doi.org/10.1038/ng.864>.
- [53] L. Gahurova, S. ichi Tomizawa, S.A. Smallwood, K.R. Stewart-Morgan, H. Saadeh, J. Kim, S.R. Andrews, T. Chen, G. Kelsey, Transcription and chromatin determinants of de novo DNA methylation timing in oocytes, *Epigenetics and Chromatin* 10 (2017) 25, <https://doi.org/10.1186/s13072-017-0133-5>.
- [54] B. Zhang, H. Zheng, B. Huang, W. Li, Y. Xiang, X. Peng, J. Ming, X. Wu, Y. Zhang, Q. Xu, W. Liu, X. Kou, Y. Zhao, W. He, C. Li, B. Chen, Y. Li, Q. Wang, J. Ma, Q. Yin, K. Ke, A. Meng, S. Gao, F. Xu, J. Na, W. Xie, Allelic reprogramming of the histone modification H3K4me3 in early mammalian development, *Nature* 537 (2016) 553–557, <https://doi.org/10.1038/nature19361>.
- [55] J.A. Dahl, I. Jung, H. Aanes, G.D. Greggains, A. Manaf, M. Lerdrup, G. Li, S. Kuan, B. Li, A.Y. Lee, S. Preissl, I. Jermstad, M.H. Haugen, R. Suganthan, M. Björås, K. Hansen, K.T. Dalen, P. Fedorcsak, B. Ren, A. Klungland, Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition, *Nature* 537 (2016) 548–552, <https://doi.org/10.1038/nature19360>.
- [56] C.W. Hanna, A. Taudt, J. Huang, L. Gahurova, A. Kranz, S. Andrews, W. Dean, A.F. Stewart, M. Colomé-Tatché, G. Kelsey, MLL2 conveys transcription-independent H3K4 trimethylation in oocytes, *Nat. Struct. Mol. Biol.* 25 (2018) 73–82, <https://doi.org/10.1038/s41594-017-0013-5>.
- [57] C.V. Andreu-Vieyra, R. Chen, J.E. Agno, S. Glaser, K. Anastasiadis, A. Stewart Francis, M.M. Matzuk, MLL2 is required in oocytes for bulk histone 3 lysine 4 trimethylation and transcriptional silencing, *PLoS Biol.* 8 (2010) 53–54, <https://doi.org/10.1371/journal.pbio.1000453>.
- [58] K.S. Voo, D.L. Carlone, B.M. Jacobsen, A. Flodin, D.G. Skalknik, Cloning of a mammalian transcriptional activator that binds unmethylated CpG motifs and shares a CXXC domain with DNA methyltransferase, human trithorax, and methyl-CpG binding domain protein 1, *Mol. Cell Biol.* 20 (2000) 2108–2121, <https://doi.org/10.1128/MCB.20.6.2108-2121.2000>.
- [59] D.L. Carlone, D.G. Skalknik, CpG binding protein is crucial for early embryonic development, *Mol. Cell Biol.* 21 (2001) 7601–7606, <https://doi.org/10.1128/mcb.21.22.7601-7606.2001>.
- [60] C. Yu, X. Fan, Q.Q. Sha, H.H. Wang, B.T. Li, X.X. Dai, L. Shen, J. Liu, L. Wang, K. Liu, F. Tang, H.Y. Fan, CFP1 regulates histone H3K4 trimethylation and developmental potential in mouse oocytes, *Cell Rep.* 20 (2017) 1161–1172, <https://doi.org/10.1016/j.celrep.2017.07.011>.
- [61] D.A. Brown, V. Di Cerbo, A. Feldmann, J. Ahn, S. Ito, N.P. Blackledge, M. Nakayama, M. McClellan, E. Dimitrova, A.H. Turberfield, H.K. Long, H.W. King, S. Kriacounis, L. Schermelleh, T.G. Kutateladze, H. Koseki, R.J. Klose, The SET1 complex selects actively transcribed target genes via multivalent interaction with CpG island chromatin, *Cell Rep.* 20 (2017) 2313–2327, <https://doi.org/10.1016/j.celrep.2017.08.030>.
- [62] J.H. Lee, K.S. Voo, D.G. Skalknik, Identification and characterization of the DNA binding domain of CpG-binding protein, *J. Biol. Chem.* 276 (2001) 44669–44676, <https://doi.org/10.1074/jbc.M107192000>.
- [63] C. Xu, C. Bian, R. Lam, A. Dong, J. Min, The structural basis for selective binding of non-methylated CpG islands by the CFP1 CXXC domain, *Nat. Commun.* 2 (2011) 227, <https://doi.org/10.1038/ncomms1237>.
- [64] V. Krauss, G. Reuter, DNA methylation in drosophila—a critical evaluation, *Prog. Mol. Biol. Transl. Sci.* Elsevier B.V., 2011, pp. 177–191, <https://doi.org/10.1016/B978-0-12-387685-0.00003-2>.

- [65] L.M. Iyer, S. Abhiman, L. Aravind, Natural history of eukaryotic DNA methylation systems, *Prog. Mol. Biol. Transl. Sci.* Elsevier B.V., 2011, pp. 25–104, <https://doi.org/10.1016/B978-0-12-387685-0.00002-0>.
- [66] M. Birke, S. Schreiner, M.-P. García-Cuellar, K. Mahr, F. Titgemeyer, R.K. Slany, The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation, *Nucleic Acids Res.* 30 (2002) 958–965, <https://doi.org/10.1093/nar/30.4.958>.
- [67] C. Xu, K. Liu, M. Lei, A. Yang, Y. Li, T.R. Hughes, J. Min, DNA sequence recognition of human CXXC domains and their structural determinants, *Structure* 26 (2018) 85–95.e3, <https://doi.org/10.1016/j.str.2017.11.022>.
- [68] M.D. Allen, C.G. Grummitt, C. Hilcenko, S.Y. Min, L.M. Tonkin, C.M. Johnson, S.M. Freund, M. Bycroft, A.J. Warren, Solution structure of the nonmethyl-CpG-binding CXXC domain of the leukaemia-associated MLL histone methyltransferase, *EMBO J.* 25 (2006) 4503–4512, <https://doi.org/10.1038/sj.emboj.7601340>.
- [69] T. Cierpicki, L.E. Risner, J. Grembecka, S.M. Lukasik, R. Popovic, M. Omonkowska, D.D. Shultis, N.J. Zeleznik-Le, J.H. Bushweller, Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia, *Nat. Struct. Mol. Biol.* 17 (2010) 62–69, <https://doi.org/10.1038/nsmb.1714>.
- [70] C. Bach, D. Mueller, S. Buhl, M.P. García-Cuellar, R.K. Slany, Alterations of the CxxC domain preclude oncogenic activation of mixed-lineage leukemia 2, *Oncogene* 28 (2009) 815–823, <https://doi.org/10.1038/ncr.2008.443>.
- [71] T.A. Milne, J. Kim, G.G. Wang, S.C. Stadler, V. Basur, S.J. Whitcomb, Z. Wang, A.J. Ruthenburg, K.S.J. Elenitoba-Johnson, R.G. Roeder, C.D. Allis, Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis, *Mol. Cell* 38 (2010) 853–863, <https://doi.org/10.1016/j.molcel.2010.05.011>.
- [72] D. Hu, X. Gao, K. Cao, M.A. Morgan, G. Mas, E.R. Smith, A.G. Volk, E.T. Bartom, J.D. Crispino, L. Di Croce, A. Shilatfard, Not all H3K4 methylations are created equal: MLL2/COMPASS dependency in primordial germ cell specification, *Mol. Cell* 65 (2017) 460–475.e6, <https://doi.org/10.1016/j.molcel.2017.01.013>.
- [73] C. Meyer, T. Burmeister, D. Gröger, G. Tsaar, L. Fechina, A. Renneville, R. Sutton, N.C. Venn, M. Emerenciano, M.S. Pombo-De-Oliveira, C. Barbieri Blunck, B. Almeida Lopes, J. Zuna, J. Trka, P. Ballerini, H. Lapillonne, M. De Brackeleer, G. Cazzaniga, L. Corral Abascal, V.H.J. Van Der Velden, E. Delabesse, T.S. Park, S.H. Oh, M.L.M. Silva, T. Lund-Aho, V. Juvonen, A.S. Moore, O. Heidenreich, J. Vormoor, E. Zerkalenkova, Y. Olshanskaya, C. Bueno, P. Menendez, A. Teigler-Schlegel, U. Zur Stadt, J. Lentes, G. Göhring, A. Kustanovich, O. Aleinikova, B.W. Schäfer, S. Kubetzko, H.O. Madsen, B. Gruhn, X. Duarte, P. Gameiro, E. Lippert, A. Bidet, J.M. Cayuela, E. Clappier, C.N. Alonso, C.M. Zwaan, M.M. Van Den Heuvel-Eibrink, S. Izraeli, L. Trakhtenbrot, P. Archer, J. Hancock, A. Möricke, J. Alten, M. Schrappe, M. Stanulla, S. Strehl, A. Attarbaschi, M. Dworzak, O.A. Haas, R. Panzer-Grümayer, L. Sedék, T. Szczepa, A. Caye, L. Suarez, H. Cavé, R. Marschalek, The MLL recombinome of acute leukemias in 2017, *Leukemia* 32 (2018) 273–284, <https://doi.org/10.1038/leu.2017.213>.
- [74] N.J. Zeleznik-Le, A.M. Harden, J.D. Rowley, 11q23 translocations split the “AT-hook” cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene, *Proc. Natl. Acad. Sci. U. S. A.* 91 (1994) 10610–10614, <https://doi.org/10.1073/pnas.91.22.10610>.
- [75] A.C. Winters, K.M. Bernt, MLL-rearranged leukemias- an update on science and clinical approaches, *Front. Pediatr.* 5 (2017), <https://doi.org/10.3389/fped.2017.00004>.
- [76] P.M. Ayton, M.L. Cleary, Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins, *Oncogene* 20 (2001) 5695–5707, <https://doi.org/10.1038/sj.onc.1204639>.
- [77] A.G. Muntean, J.L. Hess, The pathogenesis of mixed-lineage leukemia, *Annu. Rev. Pathol. Mech. Dis.* 7 (2012) 283–301, <https://doi.org/10.1146/annurev-pathol-011811-132434>.
- [78] P.M. Ayton, E.H. Chen, M.L. Cleary, Binding to nonmethylated CpG DNA is essential for target recognition, transactivation, and myeloid transformation by an MLL oncoprotein, *Mol. Cell Biol.* 24 (2004) 10470–10478, <https://doi.org/10.1128/mcb.24.23.10470-10478.2004>.
- [79] L.E. Risner, A. Kuntimaddi, A.A. Lokken, N.J. Achille, N.W. Birch, K. Schoenfeldt, J.H. Bushweller, N.J. Zeleznik-Le, Functional specificity of CpG DNA-binding CXXC domains in mixed lineage leukemia, *J. Biol. Chem.* 288 (2013) 29901–29910, <https://doi.org/10.1074/jbc.M113.474858>.
- [80] T. Rozovskaia, O. Ravid-Amir, S. Tillib, G. Getz, E. Feinstein, H. Agrawal, A. Nagler, E.F. Rappaport, I. Issaeva, Y. Matsuo, U.R. Kees, T. Lapidot, F. Lo Coco, R. Foa, A. Mazo, T. Nakamura, C.M. Croce, G. Cimino, E. Domany, E. Canaani, Expression profiles of acute lymphoblastic and myeloblastic leukemias with ALL-1 rearrangements, *Proc. Natl. Acad. Sci. U. S. A.* 100 (2003) 7853–7858, <https://doi.org/10.1073/pnas.1132115100>.
- [81] S.A. Armstrong, J.E. Staunton, L.B. Silverman, R. Pieters, M.L. Den Boer, M.D. Minden, S.E. Sallan, E.S. Lander, T.R. Golub, S.J. Korsmeyer, MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia, *Nat. Genet.* 30 (2002) 41–47, <https://doi.org/10.1038/ng765>.
- [82] J. Kerry, L. Godfrey, E. Repapi, M. Tapia, N.P. Blackledge, H. Ma, E. Ballabio, S. O’Byrne, F. Ponthan, O. Heidenreich, A. Roy, I. Roberts, M. Konopleva, R.J. Klose, H. Geng, T.A. Milne, MLL-AF4 spreading identifies binding sites that are distinct from super-enhancers and that govern sensitivity to DOT1L inhibition in leukemia, *Cell Rep.* 18 (2017) 482–495, <https://doi.org/10.1016/j.celrep.2016.12.054>.
- [83] C.C. Sze, K. Cao, C.K. Collings, S.A. Marshall, E.J. Rendleman, P.A. Ozark, F.X. Chen, M.A. Morgan, L. Wang, A. Shilatfard, Histone H3K4 methylation-dependent and -independent functions of set1A/COMPASS in embryonic stem cell self-renewal and differentiation, *Genes Dev.* 31 (2017) 1732–1737, <https://doi.org/10.1101/gad.303768.117>.
- [84] A.J. Ruthenburg, H. Li, D.J. Patel, C. David Allis, Multivalent engagement of chromatin modifications by linked binding modules, *Nat. Rev. Mol. Cell Biol.* 8 (2007) 983–994, <https://doi.org/10.1038/nrm2298>.
- [85] J. Mahadevan, D.G. Skalik, Efficient differentiation of murine embryonic stem cells requires the binding of CXXC finger protein 1 to DNA or methylated histone H3-Lys4, *Gene* 594 (2016) 1–9, <https://doi.org/10.1016/j.gene.2016.08.048>.
- [86] H.C. Eberl, C.G. Spruijt, C.D. Kelstrup, M. Vermeulen, M. Mann, A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics, *Mol. Cell* 49 (2013) 368–378, <https://doi.org/10.1016/j.molcel.2012.10.026>.
- [87] Z. Wang, J. Song, T.A. Milne, G.G. Wang, H. Li, C.D. Allis, D.J. Patel, Pro isomerization in MLL1 PHD3-Bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression, *Cell* 141 (2010) 1183–1194, <https://doi.org/10.1016/j.cell.2010.05.016>.
- [88] P.Y. Chang, R.A. Hom, C.A. Musselman, L. Zhu, A. Kuo, O. Gozani, T.G. Kutateladze, M.L. Cleary, Binding of the MLL PHD3 finger to histone H3K4me3 is required for MLL-dependent gene transcription, *J. Mol. Biol.* 400 (2010) 137–144, <https://doi.org/10.1016/j.jmb.2010.05.005>.
- [89] M. Vermeulen, H.C. Eberl, F. Matarese, H. Marks, S. Denissov, F. Butter, K.K. Lee, J.V. Olsen, A.A. Hyman, H.G. Stunnenberg, M. Mann, Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers, *Cell* 142 (2010) 967–980, <https://doi.org/10.1016/j.cell.2010.08.020>.
- [90] S. Park, U. Osmers, G. Raman, R.H. Schwantes, M.O. Diaz, J.H. Bushweller, The PHD3 domain of MLL acts as a CYP33-regulated switch between MLL-mediated activation and repression, *Biochemistry* 49 (2010) 6576–6586, <https://doi.org/10.1021/bi1009387>.
- [91] M. Wu, P.F. Wang, J.S. Lee, S. Martin-Brown, L. Florens, M. Washburn, A. Shilatfard, Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS, *Mol. Cell Biol.* 28 (2008) 7337–7344, <https://doi.org/10.1128/mcb.00976-08>.
- [92] M.B. Ardehali, A. Mei, K.L. Zobeck, M. Caron, J.T. Lis, T. Kusch, Drosophila Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription, *EMBO J.* 30 (2011) 2817–2828, <https://doi.org/10.1038/emboj.2011.194>.
- [93] G. Hallson, R.E. Hollebakk, T. Li, M. Syrzycka, I. Kim, S. Cotsworth, K.A. Fitzpatrick, D.A.R. Sinclair, B.M. Honda, dSet1 is the main H3K4 Di- and trimethyltransferase throughout Drosophila development, *Genetics* 190 (2012) 91–100, <https://doi.org/10.1534/genetics.111.135863>.
- [94] L.M.I. Austenaa, I. Barozzi, M. Simonatto, S. Masella, G. Della Chiara, S. Ghisletti, A. Curina, E. de Wit, B.A.M. Bouwman, S. de Pretis, V. Piccolo, A. Termanini, E. Prosperini, M. Pelizzola, W. de Laat, G. Natoli, Transcription of mammalian cis-regulatory elements is restrained by actively enforced early termination, *Mol. Cell* 60 (2015) 460–474, <https://doi.org/10.1016/j.molcel.2015.09.018>.
- [95] C.C. Ebmeier, B. Erickson, B.L. Allen, M.A. Allen, H. Kim, N. Fong, J.R. Jacobsen, K. Liang, A. Shilatfard, R.D. Dowell, W.M. Old, D.L. Bentley, D.J. Taatjes, Human TFIIF kinase CDK7 regulates transcription-associated chromatin modifications, *Cell Rep.* 20 (2017) 1173–1186, <https://doi.org/10.1016/j.celrep.2017.07.021>.
- [96] J.-H. Lee, D.G. Skalik, Wdr82 is a C-terminal domain-binding protein that recruits the Set1A histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes, *Mol. Cell Biol.* 28 (2008) 609–618, <https://doi.org/10.1128/mcb.01356-07>.
- [97] A.G. Muntean, J. Tan, K. Sitwala, Y. Huang, J. Bronstein, J.A. Connelly, V. Basur, K.S.J. Elenitoba-Johnson, J.L. Hess, The PAF complex synergizes with MLL fusion proteins at HOX loci to promote leukemogenesis, *Cancer Cell* 17 (2010) 609–621, <https://doi.org/10.1016/j.ccr.2010.04.012>.
- [98] L. Hou, Y. Wang, Y. Liu, N. Zhang, I. Shamovsky, E. Nudler, B. Tian, B.D. Dynlacht, Paf1C regulates RNA polymerase II progression by modulating elongation rate, *Proc. Natl. Acad. Sci. U. S. A.* 116 (2019) 14583–14592, <https://doi.org/10.1073/pnas.1904324116>.
- [99] J. Kim, J.A. Kim, R.K. McGinty, U.T.T. Nguyen, T.W. Muir, C.D. Allis, R.G. Roeder, The n-SET domain of Set1 regulates H2B ubiquitylation-dependent H3K4 methylation, *Mol. Cell* 49 (2013) 1121–1133, <https://doi.org/10.1016/j.molcel.2013.01.034>.
- [100] H. Xue, T. Yao, M. Cao, G. Zhu, Y. Li, G. Yuan, Y. Chen, M. Lei, J. Huang, Structural basis of nucleosome recognition and modification by MLL methyltransferases, *Nature* 573 (2019) 445–449, <https://doi.org/10.1038/s41586-019-1528-1>.
- [101] R. Pavri, B. Zhu, G. Li, P. Trojer, S. Mandal, A. Shilatfard, D. Reinberg, Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II, *Cell* 125 (2006) 703–717, <https://doi.org/10.1016/j.cell.2006.04.029>.
- [102] J. Kim, M. Guermah, R.K. McGinty, J.S. Lee, Z. Tang, T.A. Milne, A. Shilatfard, T.W. Muir, R.G. Roeder, RAD6-mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells, *Cell* 137 (2009) 459–471, <https://doi.org/10.1016/j.cell.2009.02.027>.
- [103] P.L. Hsu, H. Shi, C. Leonen, J. Kang, C. Chatterjee, N. Zheng, Structural basis of H2B ubiquitylation-dependent H3K4 methylation by COMPASS, *Mol. Cell* 76 (2019) 712–723.e4, <https://doi.org/10.1016/j.molcel.2019.10.013>.
- [104] E.J. Worden, X. Zhang, C. Wolberger, Structural basis for COMPASS recognition of an H2B-ubiquitinated nucleosome, *Elife* 9 (2020), <https://doi.org/10.7554/eLife.53199>.
- [105] W. Xie, S. Nagarajan, S.J. Baumgart, R.L. Kosinsky, Z. Najafova, V. Kari, M. Hennion, D. Indenbirken, S. Bonn, A. Grundhoff, F. Wegwitz, A. Mansouri, S.A. Johnsen, RNF40 regulates gene expression in an epigenetic context-dependent manner, *Genome Biol.* 18 (2017) 32, <https://doi.org/10.1186/s13059-017-1159-5>.



- [106] I. Jung, S.K. Kim, M. Kim, Y.M. Han, Y.S. Kim, D. Kim, D. Lee, H2B mono-ubiquitylation is a 5'-enriched active transcription mark and correlates with exon-intron structure in human cells, *Genome Res.* 22 (2012) 1026–1035, <https://doi.org/10.1101/gr.120634.111>.
- [107] V. Vethantham, Y. Yang, C. Bowman, P. Asp, J.-H. Lee, D.G. Skalnik, B.D. Dynlacht, Dynamic loss of H2B ubiquitylation without corresponding changes in H3K4 trimethylation during myogenic differentiation, *Mol. Cell. Biol.* 32 (2012) 1044–1055, <https://doi.org/10.1128/mcb.06026-11>.
- [108] Y.H. Lee, H. Ma, T.Z. Tan, S.S. Ng, R. Soong, S. Mori, X.Y. Fu, M. Zernicka-Goetz, Q. Wu, Protein arginine methyltransferase 6 regulates embryonic stem cell identity, *Stem Cells Dev.* 21 (2012) 2613–2622, <https://doi.org/10.1089/scd.2011.0330>.
- [109] C. Bouchard, P. Sahu, M. Meixner, R.R. Nötzold, M.B. Rust, E. Kremmer, R. Feederle, G. Hart-Smith, F. Finkernagel, M. Bartkuhn, S. Savai Pullamsetti, A. Nist, T. Stiewe, S. Philipsen, U.M. Bauer, Genomic location of PRMT6-dependent H3R2 methylation is linked to the transcriptional outcome of associated genes, *Cell Rep.* 24 (2018) 3339–3352, <https://doi.org/10.1016/j.celrep.2018.08.052>.
- [110] E. Guccione, C. Bassi, F. Casadio, F. Martinato, M. Cesaroni, H. Schuchlaltz, B. Lüscher, B. Amati, Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive, *Nature* 449 (2007) 933–937, <https://doi.org/10.1038/nature06166>.
- [111] D. Hyllus, C. Stein, K. Schnabel, E. Schiltz, A. Imhof, Y. Dou, J. Hsieh, U.M. Bauer, PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation, *Genes Dev.* 21 (2007) 3369–3380, <https://doi.org/10.1101/gad.447007>.
- [112] Y. Li, H. Li, Many keys to push: diversifying the “readership” of plant homeodomain fingers, *Acta Biochim. Biophys. Sin. Shanghai* 44 (2012) 28–39, <https://doi.org/10.1093/abbs/gmr117>.
- [113] C. He, N. Liu, D. Xie, Y. Liu, Y. Xiao, F. Li, Structural basis for histone H3K4me3 recognition by the N-terminal domain of the PHD finger protein Spp1, *Biochem. J.* 476 (2019) 1957–1973, <https://doi.org/10.1042/BCJ20190091>.
- [114] M. Vermeulen, K.W. Mulder, S. Denissov, W.W.M.P. Pijnappel, F.M.A. van Schaik, R.A. Varier, M.P.A. Baltissen, H.G. Stunnenberg, M. Mann, H.T.M. Timmers, Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4, *Cell* 131 (2007) 58–69, <https://doi.org/10.1016/j.cell.2007.08.016>.
- [115] H. van Ingen, F.M.A. van Schaik, H. Wienk, J. Ballering, H. Rehmann, A.C. Dechesne, J.A.W. Kruijzer, R.M.J. Liskamp, H.T.M. Timmers, R. Boelens, Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3, *Structure* 16 (2008) 1245–1256, <https://doi.org/10.1016/j.str.2008.04.015>.
- [116] A.N. Iberg, A. Espejo, D. Cheng, D. Kim, J. Michaud-Levesque, S. Richard, M.T. Bedford, Arginine methylation of the histone H3 tail impedes effector binding, *J. Biol. Chem.* 283 (2008) 3006–3010, <https://doi.org/10.1074/jbc.C700192200>.
- [117] R.S.L. Stein, W. Wang, The recognition specificity of the CHD1 chromodomain with modified histone H3 peptides, *J. Mol. Biol.* 406 (2011) 527–541, <https://doi.org/10.1016/j.jmb.2010.12.030>.
- [118] Y. Zheng, J.D. Tipton, P.M. Thomas, N.L. Kelleher, S.M.M. Sweet, Site-specific human histone H3 methylation stability: fast K4me3 turnover, *Proteomics* 14 (2014) 2190–2199, <https://doi.org/10.1002/prot.201400060>.
- [119] J. Christensen, K. Agger, P.A.C. Cloos, D. Pasini, S. Rose, L. Sennels, J. Rappsilber, K.H. Hansen, A.E. Salcini, K. Helin, RBP2 belongs to a family of demethylases, specific for tri- and dimethylated lysine 4 on histone 3, *Cell* 128 (2007) 1063–1076, <https://doi.org/10.1016/j.cell.2007.02.003>.
- [120] S. Iwase, F. Lan, P. Bayliss, L. de la Torre-Ubieta, M. Huarte, H.H. Qi, J.R.R. Whetstone, A. Bonni, T.M. Roberts, Y. Shi, The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases, *Cell* 128 (2007) 1077–1088, <https://doi.org/10.1016/j.cell.2007.02.017>.
- [121] R.J. Klose, Q. Yan, Z. Tothova, K. Yamane, H. Erdjument-Bromage, P. Tempst, D.G. Gilliland, Y. Zhang, W.G. Kaelin, The retinoblastoma binding protein RBP2 is an H3K4 demethylase, *Cell* 128 (2007) 889–900, <https://doi.org/10.1016/j.cell.2007.02.013>.
- [122] K. Yamane, K. Tateishi, R.J. Klose, J. Fang, L.A. Fabrizio, H. Erdjument-Bromage, J. Taylor-Papadimitriou, P. Tempst, Y. Zhang, PLU-1 is an H3K4 demethylase involved in transcriptional repression and breast cancer cell proliferation, *Mol. Cell* 25 (2007) 801–812, <https://doi.org/10.1016/j.molcel.2007.03.001>.
- [123] M.G. Lee, J. Norman, A. Shilatfard, R. Shiekhata, Physical and functional association of a trimethyl H3K4 demethylase and Ring6a/MBL, a polycomb-like protein, *Cell* 128 (2007) 877–887, <https://doi.org/10.1016/j.cell.2007.02.004>.
- [124] B. Habiha, B.L. Kidder, KDM5B is a master regulator of the H3K4-methylome in stem cells, development and cancer, *Semin. Cancer Biol.* 57 (2019) 79–85, <https://doi.org/10.1016/j.semcancer.2018.11.001>.
- [125] P.A.C. Cloos, J. Christensen, K. Agger, K. Helin, Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease, *Genes Dev.* 22 (2008) 1115–1140, <https://doi.org/10.1101/gad.1652908>.
- [126] S.M. Kooistra, K. Helin, Post-translational modifications: molecular mechanisms and potential functions of histone demethylases, *Nat. Rev. Mol. Cell Biol.* 13 (2012) 297–311, <https://doi.org/10.1038/nrm3327>.
- [127] A. Patsialou, D. Wilsker, E. Moran, DNA-binding properties of ARID family proteins, *Nucleic Acids Res.* 33 (2005) 66–80, <https://doi.org/10.1093/nar/gki145>.
- [128] R.J. Klose, E.M. Kallin, Y. Zhang, JmjC-domain-containing proteins and histone demethylation, *Nat. Rev. Genet.* 7 (2006) 715–727, <https://doi.org/10.1038/nrg1945>.
- [129] G.G. Wang, J. Song, Z. Wang, H.L. Dormann, F. Casadio, H. Li, J.L. Luo, D.J. Patel, C.D. Allis, Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger, *Nature* 459 (2009) 847–851, <https://doi.org/10.1038/nature08036>.
- [130] B.L. Kidder, G. Hu, K. Zhao, KDM5B focuses H3K4 methylation near promoters and enhancers during embryonic stem cell self-renewal and differentiation, *Genome Biol.* 15 (2014) R32, <https://doi.org/10.1186/gb-2014-15-2-r32>.
- [131] L. Xie, C. Pelz, W. Wang, A. Bashar, O. Varlamova, S. Shadle, S. Impey, KDM5B regulates embryonic stem cell self-renewal and represses cryptic intragenic transcription, *EMBO J.* 30 (2011) 1473–1484, <https://doi.org/10.1038/emboj.2011.91>.
- [132] J.E. Longbotham, C.M. Chio, V. Dharmarajan, M.J. Trnka, I.O. Torres, D. Goswami, K. Ruiz, A.L. Burlingame, P.R. Griffin, D.G. Fujimori, Histone H3 binding to the PHD1 domain of histone demethylase KDM5A enables active site remodeling, *Nat. Commun.* 10 (2019) 94, <https://doi.org/10.1038/s41467-018-07829-z>.
- [133] I.O. Torres, K.M. Kuchenbecker, C.I. Nnadi, R.J. Fletterick, M.J.S. Kelly, D.G. Fujimori, Histone demethylase KDM5A is regulated by its reader domain through a positive-feedback mechanism, *Nat. Commun.* 6 (2015) 6204, <https://doi.org/10.1038/ncomms7204>.
- [134] Y. Zhang, H. Yang, X. Guo, N. Rong, Y. Song, Y. Xu, W. Lan, X. Zhang, M. Liu, Y. Xu, C. Cao, The PHD1 finger of KDM5B recognizes unmodified H3K4 during the demethylation of histone H3K4me2/3 by KDM5B, *Protein Cell* 5 (2014) 837–850, <https://doi.org/10.1007/s13238-014-0078-4>.
- [135] M. Albert, S.U. Schmitz, S.M. Kooistra, M. Malatesta, C. Morales Torres, J.C. Reikling, J.V. Johansen, I. Abarrategui, K. Helin, The histone demethylase Jarid1b ensures faithful mouse development by protecting developmental genes from aberrant H3K4me3, *PLoS Genet.* 9 (2013) e1003461, <https://doi.org/10.1371/journal.pgen.1003461>.
- [136] D. Sproul, R.R. Meehan, Genomic insights into cancer-associated aberrant CpG island hypermethylation, *Brief. Funct. Genomics.* 12 (2013) 174–190, <https://doi.org/10.1093/bfpg/els063>.
- [137] M. Klutstein, D. Nejman, R. Greenfield, H. Cedar, DNA methylation in cancer and aging, *Cancer Res.* 76 (2016) 3446–3450, <https://doi.org/10.1158/0008-5472.CAN-15-3278>.
- [138] A.M. Deaton, A. Bird, CpG islands and the regulation of transcription, *Genes Dev.* 25 (2011) 1010–1022, <https://doi.org/10.1101/gad.2037511>.
- [139] H.K. Long, N.P. Blackledge, R.J. Klose, ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection, *Biochem. Soc. Trans.* 41 (2013) 727–740, <https://doi.org/10.1042/BST20130028>.
- [140] A. Meissner, T.S. Mikkelsen, H. Gu, M. Wernig, J. Hanna, A. Sivachenko, X. Zhang, B.E. Bernstein, C. Nusbaum, D.B. Jaffe, A. Gnirke, R. Jaenisch, E.S. Lander, Genome-scale DNA methylation maps of pluripotent and differentiated cells, *Nature* 454 (2008) 766–770, <https://doi.org/10.1038/nature07107>.
- [141] M. Morselli, W.A. Pastor, B. Montanini, K. Nee, R. Ferrari, K. Fu, G. Bonora, L. Rubbi, A.T. Clark, S. Ottonello, S.E. Jacobsen, M. Pellegrini, In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse, *Elife* 2015 (2015) e06205, <https://doi.org/10.7554/eLife.06205>.
- [142] P. Singh, A.X. Li, D.A. Tran, N. Oates, E.R. Kang, X. Wu, P.E. Szabo, De novo DNA methylation in the male germ line occurs by default but is excluded at sites of H3K4 methylation, *Cell Rep.* 4 (2013) 205–219, <https://doi.org/10.1016/j.celrep.2013.06.004>.
- [143] M. Weber, I. Hellmann, M.B. Stadler, L. Ramos, S. Pääbo, M. Rebhan, D. Schübeler, Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome, *Nat. Genet.* 39 (2007) 457–466, <https://doi.org/10.1038/ng1990>.
- [144] D.N. Ciccone, H. Su, S. Hevi, F. Gay, H. Lei, J. Bajko, G. Xu, E. Li, T. Chen, KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints, *Nature* 461 (2009) 415–418, <https://doi.org/10.1038/nature08315>.
- [145] K.R. Stewart, L. Veselovska, J. Kim, J. Huang, H. Saadeh, S.I. Tomizawa, S.A. Smallwood, T. Chen, G. Kelsey, Dynamic changes in histone modifications precede de novo DNA methylation in oocytes, *Genes Dev.* 29 (2015) 2449–2462, <https://doi.org/10.1101/gad.271353.115>.
- [146] J. Wang, S. Hevi, J.K. Kurash, H. Lei, F. Gay, J. Bajko, H. Su, W. Sun, H. Chang, G. Xu, F. Gaudet, E. Li, T. Chen, The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation, *Nat. Genet.* 41 (2009) 125–129, <https://doi.org/10.1038/ng.268>.
- [147] R. Terranova, H. Agherbi, A. Boned, S. Meresse, M. Djabali, Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll, *Proc. Natl. Acad. Sci.* 103 (2006) 6629–6634, <https://doi.org/10.1073/pnas.0507425103>.
- [148] J.L. Hu, B.O. Zhou, R.R. Zhang, K.L. Zhang, J.Q. Zhou, G.L. Xu, The N-terminus of histone H3 is required for de novo DNA methylation in chromatin, *Proc. Natl. Acad. Sci. U. S. A.* 106 (2009) 22187–22192, <https://doi.org/10.1073/pnas.0905767106>.
- [149] H. Gowher, A. Jeltsch, Mammalian DNA methyltransferases: new discoveries and open questions, *Biochem. Soc. Trans.* 46 (2018) 1191–1202, <https://doi.org/10.1042/BST20170574>.
- [150] S.K.T. Ooi, C. Qiu, E. Bernstein, K. Li, D. Jia, Z. Yang, H. Erdjument-Bromage, P. Tempst, S.P. Lin, C.D. Allis, X. Cheng, T.H. Bestor, DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA, *Nature* 448 (2007) 714–717, <https://doi.org/10.1038/nature05987>.
- [151] J. Otani, T. Nankumo, K. Arita, S. Inamoto, M. Ariyoshi, M. Shirakawa, Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain, *EMBO Rep.* 10 (2009) 1235–1241, <https://doi.org/10.1038/embo.2009.218>.
- [152] Y. Zhang, R. Jurkowska, S. Soerens, A. Rajavelu, A. Dhayan, I. Bock, P. Rathert, O. Brandt, R. Reinhardt, W. Fischle, A. Jeltsch, Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the



- histone H3 tail, *Nucleic Acids Res.* 38 (2010) 4246–4253, <https://doi.org/10.1093/nar/gkq147>.
- [153] K.M. Noh, H. Wang, H.R. Kim, W. Wenderski, F. Fang, C.H. Li, S. Dewell, S.H. Hughes, A.M. Melnick, D.J. Patel, H. Li, C.D. Allis, Engineering of a histone-recognition domain in Dnmt3a alters the epigenetic landscape and phenotypic features of mouse ESCs, *Mol. Cell* 59 (2015) 89–103, <https://doi.org/10.1016/j.molcel.2015.05.017>.
- [154] T. Chen, N. Tsujimoto, E. Li, The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin, *Mol. Cell Biol.* 24 (2004) 9048–9058, <https://doi.org/10.1128/mcb.24.20.9048-9058.2004>.
- [155] Y.Z. Ge, M.T. Pu, H. Gowher, H.P. Wu, J.P. Ding, A. Jeltsch, G.L. Xu, Chromatin targeting of de novo DNA methyltransferases by the PWWP domain, *J. Biol. Chem.* 279 (2004) 25447–25454, <https://doi.org/10.1074/jbc.M312296200>.
- [156] B.Z. Li, Z. Huang, Q.Y. Cui, X.H. Song, L. Du, A. Jeltsch, P. Chen, G. Li, E. Li, G.L. Xu, Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase, *Cell Res.* 21 (2011) 1172–1181, <https://doi.org/10.1038/cr.2011.92>.
- [157] X. Guo, L. Wang, J. Li, Z. Ding, J. Xiao, X. Yin, S. He, P. Shi, L. Dong, G. Li, C. Tian, J. Wang, Y. Cong, Y. Xu, Structural insight into autoinhibition and histone H3-induced activation of DNMT3A, *Nature* 517 (2015) 640–644, <https://doi.org/10.1038/nature13899>.
- [158] W. Feng, M. Yonezawa, J. Ye, T. Jenuwein, I. Grummt, PHF8 activates transcription of rRNA genes through H3K4me3 binding and H3K9me1/2 demethylation, *Nat. Struct. Mol. Biol.* 17 (2010) 445–450, <https://doi.org/10.1038/nsmb.1778>.
- [159] J.R. Horton, A.K. Upadhyay, H.H. Qi, X. Zhang, Y. Shi, X. Cheng, Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases, *Nat. Struct. Mol. Biol.* 17 (2010) 38–44, <https://doi.org/10.1038/nsmb.1753>.
- [160] L.R. Pack, K.R. Yamamoto, D.G. Fujimori, Opposing chromatin signals direct and regulate the activity of lysine demethylase 4C (KDM4C), *J. Biol. Chem.* 291 (2016) 6060–6070, <https://doi.org/10.1074/jbc.M115.696864>.
- [161] B. Lohse, C. Helgstrand, J.B.L. Kristensen, U. Leurs, P.A.C. Cloos, J.L. Kristensen, R.P. Clausen, Posttranslational modifications of the histone 3 tail and their impact on the activity of histone lysine demethylases in vitro, *PLoS One* 8 (2013) e67653, <https://doi.org/10.1371/journal.pone.0067653>.
- [162] H. Wu, J. Min, V.V. Lunin, T. Antoshenko, L. Dombrowski, H. Zeng, A. Allali-Hassani, V. Campagna-Slater, M. Vedadi, C.H. Arrowsmith, A.N. Plotnikov, M. Schapira, Structural biology of human H3K9 methyltransferases, *PLoS One* 5 (2010) e8570, <https://doi.org/10.1371/journal.pone.0008570>.
- [163] S. Bilodeau, M.H. Kagey, G.M. Frampton, P.B. Rahl, R.A. Young, SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state, *Genes Dev.* 23 (2009) 2484–2489, <https://doi.org/10.1101/gad.1837309>.
- [164] M. Wan, J. Liang, Y. Xiong, F. Shi, Y. Zhang, W. Lu, Q. He, D. Yang, R. Chen, D. Liu, M. Barton, Z. Songyang, The trithorax group protein Ash2 is essential for pluripotency and maintaining open chromatin in embryonic stem cells, *J. Biol. Chem.* 288 (2013) 5039–5048, <https://doi.org/10.1074/jbc.M112.424515>.
- [165] O. Bindu, G. LeRoy, D.J. Bua, B.A. Garcia, O. Gozani, S. Richard, Trimethylation of histone H3 lysine 4 impairs methylation of histone H3 lysine 9: regulation of lysine methyltransferases by physical interaction with their substrates, *Epigenetics* 5 (2010) 767–775, <https://doi.org/10.4161/epi.5.8.13278>.
- [166] H.H. Qi, M. Sarkissian, G.Q. Hu, Z. Wang, A. Bhattacharjee, D.B. Gordon, M. Gonzales, F. Lan, P.P. Ongusaba, M. Huarte, N.K. Yaghi, H. Lim, B.A. Garcia, L. Brizuela, K. Zhao, T.M. Roberts, Y. Shi, Histone H4K20/H3K9 demethylase PHF8 regulates zebrafish brain and craniofacial development, *Nature* 466 (2010) 503–507, <https://doi.org/10.1038/nature09261>.
- [167] K. Fortschegger, P. de Graaf, N.S. Outchkourov, F.M.A. van Schaik, H.T.M. Timmers, R. Shiekhattar, PHF8 targets histone methylation and RNA polymerase II to activate transcription, *Mol. Cell Biol.* 30 (2010) 3286–3298, <https://doi.org/10.1128/mcb.01520-09>.
- [168] H. Wen, J. Li, T. Song, M. Lu, P.Y. Kan, M.G. Lee, B. Sha, X. Shi, Recognition of histone H3K4 trimethylation by the plant homeodomain of PHF2 modulates histone demethylation, *J. Biol. Chem.* 285 (2010) 9322–9326, <https://doi.org/10.1074/jbc.C109.097667>.
- [169] Y. Huang, Recognition of histone H3 lysine-4 methylation by the double Tudor domain of JMJD2A, *Science* 312 (2006) 748–751, <https://doi.org/10.1126/science.1125162>.
- [170] J. Kim, J. Daniel, A. Espejo, A. Lake, M. Krishna, L. Xia, Y. Zhang, M.T. Bedford, Tudor, MBT and chromo domains gauge the degree of lysine methylation, *EMBO Rep.* 7 (2006) 397–403, <https://doi.org/10.1038/sj.embor.7400625>.
- [171] J. Lee, J.R. Thompson, M.V. Botuyan, G. Mer, Distinct binding modes specify the recognition of methylated histones H3K4 and H4K20 by JMJD2A-tudor, *Nat. Struct. Mol. Biol.* 15 (2008) 109–111, <https://doi.org/10.1038/nsmb.1326>.
- [172] K. Nishioka, S. Chuikov, K. Sarma, H. Erdjument-Bromage, C.D. Allis, P. Tempst, D. Reinberg, Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation, *Genes Dev.* 16 (2002) 479–489, <https://doi.org/10.1101/gad.967202>.
- [173] M.T. Pedersen, K. Agger, A. Laugesen, J.V. Johansen, P.A.C. Cloos, J. Christensen, K. Helin, The demethylase JMJD2C localizes to H3K4me3-positive transcription start sites and is dispensable for embryonic development, *Mol. Cell Biol.* 34 (2014) 1031–1045, <https://doi.org/10.1128/mcb.00864-13>.
- [174] P. Voigt, W.W. Tee, D. Reinberg, A double take on bivalent promoters, *Genes Dev.* 27 (2013) 1318–1338, <https://doi.org/10.1101/gad.219626.113>.
- [175] F.W. Schmitges, A.B. Prusty, M. Faty, A. Stützer, G.M. Lingaraju, J. Aiwanian, R. Lack, D. Hess, L. Li, S. Zhou, R.D. Bunker, U. Wirth, T. Bouwmeester, A. Bauer, N. Ly-Hartig, K. Zhao, H. Chan, J. Gu, H. Gut, W. Fischle, J. Müller, N.H. Thomä, Histone methylation by PRC2 is inhibited by active chromatin marks, *Mol. Cell* 42 (2011) 330–341, <https://doi.org/10.1016/j.molcel.2011.03.025>.
- [176] P. Voigt, G. LeRoy, W.J. Drury, B.M. Zee, J. Son, D.B. Beck, N.L. Young, B.A. Garcia, D. Reinberg, Asymmetrically modified nucleosomes, *Cell* 151 (2012) 181–193, <https://doi.org/10.1016/j.cell.2012.09.002>.
- [177] C.C. Lechner, N.D. Agashe, B. Fierz, Traceless synthesis of asymmetrically modified bivalent nucleosomes, *Angew. Chemie - Int. Ed.* 55 (2016) 2903–2906, <https://doi.org/10.1002/anie.201510996>.
- [178] S. Chen, L. Jiao, M. Shubbar, X. Yang, X. Liu, Unique structural platforms of Suz12 dictate distinct classes of PRC2 for chromatin binding, *Mol. Cell* 69 (2018) 840–852.e5, <https://doi.org/10.1016/j.molcel.2018.01.039>.
- [179] P. Zegerman, B. Canas, D. Pappin, T. Kouzarides, Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex, *J. Biol. Chem.* 277 (2002) 11621–11624, <https://doi.org/10.1074/jbc.C200045200>.
- [180] R.E. Mansfield, C.A. Musselman, A.H. Kwan, S.S. Oliver, A.L. Garske, F. Davrazou, J.M. Denu, T.G. Kutateladze, J.P. Mackay, Plant homeodomain (PHD) fingers of CHD4 are histone H3-binding modules with preference for unmodified H3K4 and methylated H3K9, *J. Biol. Chem.* 286 (2011) 11779–11791, <https://doi.org/10.1074/jbc.M110.208207>.
- [181] S.S. Oliver, C.A. Musselman, R. Srinivasan, J.P. Svaren, T.G. Kutateladze, J.M. Denu, Multivalent recognition of histone tails by the PHD fingers of CHD5, *Biochemistry* 51 (2012) 6534–6544, <https://doi.org/10.1021/bi3006972>.
- [182] S. Paul, A. Kuo, T. Schachl, H. Vogel, L. Joshua-Tor, W.R. McCombie, O. Gozani, M. Hammell, A.A. Mills, Chd5 requires PHD-mediated histone 3 binding for tumor suppression, *Cell Rep.* 3 (2013) 92–102, <https://doi.org/10.1016/j.celrep.2012.12.009>.
- [183] C.A. Musselman, R.E. Mansfield, A.L. Garske, F. Davrazou, A.H. Kwan, S.S. Oliver, H. O'Leary, J.M. Denu, J.P. Mackay, T.G. Kutateladze, Binding of the CHD4 PHD2 finger to histone H3 is modulated by covalent modifications, *Biochem. J.* 423 (2009) 179–187, <https://doi.org/10.1042/BJ20090870>.
- [184] C.A. Musselman, J. Ramírez, J.K. Sims, R.E. Mansfield, S.S. Oliver, J.M. Denu, J.P. Mackay, P.A. Wade, J. Hagman, T.G. Kutateladze, Bivalent recognition of nucleosomes by the tandem PHD fingers of the CHD4 ATPase is required for CHD4-mediated repression, *Proc. Natl. Acad. Sci. U. S. A.* 109 (2012) 787–792, <https://doi.org/10.1073/pnas.1113655109>.
- [185] N. Reynolds, M. Salmon-Divon, H. Dvinge, A. Hynes-Allen, G. Balasooriya, D. Leaford, A. Behrens, P. Bertone, B. Hendrich, NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression, *EMBO J.* 31 (2012) 593–605, <https://doi.org/10.1038/emboj.2011.431>.
- [186] B.E. Bernstein, T.S. Mikkelsen, X. Xie, M. Kamal, D.J. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath, R. Jaenisch, A. Wagschal, R. Feil, S.L. Schreiber, E.S. Lander, A bivalent chromatin structure marks key developmental genes in embryonic stem cells, *Cell* 125 (2006) 315–326, <https://doi.org/10.1016/j.cell.2006.02.041>.
- [187] V. Azuara, P. Perry, S. Sauer, M. Spivakov, H.F. Jørgensen, R.M. John, M. Gouti, M. Casanova, G. Warnes, M. Merkenschlager, A.G. Fisher, Chromatin signatures of pluripotent cell lines, *Nat. Cell Biol.* 8 (2006) 532–538, <https://doi.org/10.1038/ncb1403>.
- [188] T.S. Mikkelsen, M. Ku, D.B. Jaffe, B. Issac, E. Lieberman, G. Giannoukos, P. Alvarez, W. Brockman, T.K. Kim, R.P. Koche, W. Lee, E. Mendenhall, A. Oudonovan, A. Presser, C. Russ, X. Xie, A. Meissner, M. Wernig, R. Jaenisch, C. Nusbaum, E.S. Lander, B.E. Bernstein, Genome-wide maps of chromatin state in pluripotent and lineage-committed cells, *Nature* 448 (2007) 553–560, <https://doi.org/10.1038/nature06008>.
- [189] G. Mas, E. Blanco, C. Ballaré, M. Sansó, Y.G. Spill, D. Hu, Y. Aoi, F. Le Dily, A. Shilatfard, M.A. Martí-Renom, L. Di Croce, Promoter bivalency favors an open chromatin architecture in embryonic stem cells, *Nat. Genet.* 50 (2018) 1452–1462, <https://doi.org/10.1038/s41588-018-0218-5>.
- [190] M. Nikolov, A. Stützer, K. Mosch, A. Krasauskas, S. Soeroes, H. Stark, H. Urlaub, W. Fischle, Chromatin affinity purification and quantitative mass spectrometry defining the interactome of histone modification patterns, *Mol. Cell. Proteomics* 10 (2011) M110.005371, <https://doi.org/10.1074/mcp.M110.005371>.
- [191] G. Liang, J.C.Y. Lin, V. Wei, C. Yoo, J.C. Cheng, C.T. Nguyen, D.J. Weisenberger, G. Egger, D. Takai, F.A. Gonzales, P.A. Jones, Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome, *Proc. Natl. Acad. Sci. U. S. A.* 101 (2004) 7357–7362, <https://doi.org/10.1073/pnas.0401866101>.
- [192] C.E. Barnes, D.M. English, S.M. Cowley, Acetylation and Co: an expanding repertoire of histone acylations regulates chromatin and transcription, *Essays Biochem.* 63 (2019) 97–107, <https://doi.org/10.1042/EB20180061>.
- [193] E. Koutelou, C.L. Hirsch, S.Y.R. Dent, Multiple faces of the SAGA complex, *Curr. Opin. Cell Biol.* 22 (2010) 374–382, <https://doi.org/10.1016/j.celb.2010.03.005>.
- [194] C. Bian, C. Xu, J. Ruan, K.K. Lee, T.L. Burke, W. Tempel, D. Barsyte, J. Li, M. Wu, B.O. Zhou, B.E. Fleharty, A. Paulson, A. Allali-Hassani, J.Q. Zhou, G. Mer, P.A. Grant, J.L. Workman, J. Zang, J. Min, Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation, *EMBO J.* 30 (2011) 2829–2842, <https://doi.org/10.1038/emboj.2011.193>.
- [195] T. Clouaire, S. Webb, A. Bird, Cfp1 is required for gene expression-dependent H3K4 trimethylation and H3K9 acetylation in embryonic stem cells, *Genome Biol.* 15 (2014) 451, <https://doi.org/10.1186/s13059-014-0451-x>.
- [196] Z. Wang, C. Zang, K. Cui, D.E. Schones, A. Barski, W. Peng, K. Zhao, Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive

- genes, *Cell* 138 (2009) 1019–1031, <https://doi.org/10.1016/j.cell.2009.06.049>.
- [197] K. Champagne, T. Kutateladze, Structural insight into histone recognition by the ING PHD fingers, *Curr. Drug Targets* 10 (2009) 432–441, <https://doi.org/10.2174/138945009788185040>.
- [198] K.S. Champagne, N. Saksoouk, P.V. Peña, K. Johnson, M. Ullah, X.J. Yang, J. Côté, T.G. Kutateladze, The crystal structure of the ING5 PHD finger in complex with an H3K4me3 histone peptide, *Proteins Struct. Funct. Genet.* 72 (2008) 1371–1376, <https://doi.org/10.1002/prot.22140>.
- [199] S. Kim, S. Natesan, G. Cornilescu, S. Carlson, M. Tonelli, U.L. McClurg, O. Binda, C.N. Robson, J.L. Markley, S. Balaz, K.C. Glass, Mechanism of Histone H3K4me3 recognition by the plant homeodomain of inhibitor of growth 3, *J. Biol. Chem.* 291 (2016) 18326–18341, <https://doi.org/10.1074/jbc.M115.690651>.
- [200] T. Hung, O. Binda, K.S. Champagne, A.J. Kuo, K. Johnson, H.Y. Chang, M.D. Simon, T.G. Kutateladze, O. Gozani, ING4 mediates crosstalk between histone H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation, *Mol. Cell* 33 (2009) 248–256, <https://doi.org/10.1016/j.molcel.2008.12.016>.
- [201] G. Ormazza, J.A. Rodríguez, A. Ibáñez de Opakua, N. Merino, M. Villate, I. Gorroño, M. Rábano, I. Palmero, M. Vilaseca, R. Kypta, M.d.M. Vivanco, A.L. Rojas, F.J. Blanco, The tumor suppressor ING5 is a dimeric, bivalent recognition molecule of the histone H3K4me3 mark, *J. Mol. Biol.* 431 (2019) 2298–2319, <https://doi.org/10.1016/j.jmb.2019.04.018>.
- [202] P.V. Peña, F. Davrazou, X. Shi, K.L. Walter, V.V. Verkhusha, O. Gozani, R. Zhao, T.G. Kutateladze, Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2, *Nature* 442 (2006) 100–103, <https://doi.org/10.1038/nature04814>.
- [203] A. Palacios, I.G. Muñoz, D. Pantoja-Uceda, M.J. Marcaida, D. Torres, J.M. Martín-García, I. Luque, G. Montoya, F.J. Blanco, Molecular basis of histone H3K4me3 recognition by ING4, *J. Biol. Chem.* 283 (2008) 15956–15964, <https://doi.org/10.1074/jbc.M710020200>.
- [204] X. Shi, T. Hong, K.L. Walter, M. Ewalt, E. Michishita, T. Hung, D. Carney, P. Peña, F. Lan, M.R. Kaadige, N. Lacoste, C. Cayrou, F. Davrazou, A. Saha, B.R. Cairns, D.E. Ayer, T.G. Kutateladze, Y. Shi, J. Côté, K.F. Chua, O. Gozani, ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression, *Nature* 442 (2006) 96–99, <https://doi.org/10.1038/nature04835>.
- [205] P.V. Peña, R.A. Hom, T. Hung, H. Lin, A.J. Kuo, R.P.C. Wong, O.M. Subach, K.S. Champagne, R. Zhao, V.V. Verkhusha, G. Li, O. Gozani, T.G. Kutateladze, Histone H3K4me3 binding is required for the DNA repair and apoptotic activities of ING1 tumor suppressor, *J. Mol. Biol.* 380 (2008) 303–312, <https://doi.org/10.1016/j.jmb.2008.04.061>.
- [206] N. Saksoouk, N. Avvakumov, K.S. Champagne, T. Hung, Y. Doyon, C. Cayrou, E. Paquet, M. Ullah, A.J. Landry, V. Côté, X.J. Yang, O. Gozani, T.G. Kutateladze, J. Côté, HBO1 HAT complexes target chromatin throughout gene coding regions via multiple PHD finger interactions with histone H3 tail, *Mol. Cell* 33 (2009) 257–265, <https://doi.org/10.1016/j.molcel.2009.01.007>.
- [207] D. Skowrya, M. Zeremski, N. Neznanov, M. Li, Y. Choi, M. Uesugi, C.A. Hauser, W. Gu, A.V. Gudkov, J. Qin, Differential Association of Products of alternative transcripts of the candidate tumor suppressor ING1 with the mSin3/HDAC1 transcriptional corepressor complex, *J. Biol. Chem.* 276 (2001) 8734–8739, <https://doi.org/10.1074/jbc.M007664200>.
- [208] A. Kuzmichev, Y. Zhang, H. Erdjument-Bromage, P. Tempst, D. Reinberg, Role of the Sin3-histone deacetylase complex in growth regulation by the candidate tumor suppressor p33ING1, *Mol. Cell. Biol.* 22 (2002) 835–848, <https://doi.org/10.1128/mcb.22.3.835-848.2002>.
- [209] Y. Doyon, C. Cayrou, M. Ullah, A.J. Landry, V. Côté, W. Selleck, W.S. Lane, S. Tan, X.J. Yang, J. Côté, ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation, *Mol. Cell* 21 (2006) 51–64, <https://doi.org/10.1016/j.molcel.2005.12.007>.
- [210] C.A. Hazzalin, L.C. Mahadevan, Dynamic acetylation of all lysine 4-methylated histone H3 in the mouse nucleus: analysis at c-fos and c-jun, *PLoS Biol.* 3 (2005) 1–16, <https://doi.org/10.1371/journal.pbio.0030393>.
- [211] D.-W. Hsu, J.R. Chubb, T. Muramoto, C.J. Pears, L.C. Mahadevan, Dynamic acetylation of lysine-4-trimethylated histone H3 and H3 variant biology in a simple multicellular eukaryote, *Nucleic Acids Res.* 40 (2012) 7247–7256, <https://doi.org/10.1093/nar/gks367>.
- [212] N.T. Crump, C.A. Hazzalin, E.M. Bowers, R.M. Alani, P.A. Cole, L.C. Mahadevan, Dynamic acetylation of all lysine-4 trimethylated histone H3 is evolutionarily conserved and mediated by p300/CBP, *Proc. Natl. Acad. Sci. U. S. A.* 108 (2011) 7814–7819, <https://doi.org/10.1073/pnas.1100099108>.
- [213] W.K.M. Lai, B.F. Pugh, Understanding nucleosome dynamics and their links to gene expression and DNA replication, *Nat. Rev. Mol. Cell Biol.* 18 (2017) 548–562, <https://doi.org/10.1038/nrm.2017.47>.
- [214] S. Brahma, S. Henikoff, Epigenome regulation by dynamic nucleosome unwrapping, *Trends Biochem. Sci.* 45 (2020) 13–26, <https://doi.org/10.1016/j.tibs.2019.09.003>.
- [215] D.E. Schones, K. Cui, S. Cuddapah, T.Y. Roh, A. Barski, Z. Wang, G. Wei, K. Zhao, Dynamic regulation of nucleosome positioning in the human genome, *Cell* 132 (2008) 887–898, <https://doi.org/10.1016/j.cell.2008.02.022>.
- [216] A. Valouev, S.M. Johnson, S.D. Boyd, C.L. Smith, A.Z. Fire, A. Sidow, Determinants of nucleosome organization in primary human cells, *Nature* 474 (2011) 516–522, <https://doi.org/10.1038/nature10002>.
- [217] Y. Li, V.P. Schulz, C. Deng, G. Li, Y. Shen, B.K. Tusi, G. Ma, J. Stees, Y. Qiu, L.A. Steiner, L. Zhou, K. Zhao, J. Bungert, P.G. Gallagher, S. Huang, Setd1a and NURF mediate chromatin dynamics and gene regulation during erythroid lineage commitment and differentiation, *Nucleic Acids Res.* 44 (2016) 7173–7188, <https://doi.org/10.1093/nar/gkw327>.
- [218] J.T. Kurup, L.J. Campeanu, B.L. Kidder, Contribution of H3K4 demethylase KDM5B to nucleosome organization in embryonic stem cells revealed by micrococcal nuclease sequencing, *Epigenetics and Chromatin* 12 (2019) 20, <https://doi.org/10.1186/s13072-019-0266-9>.
- [219] J. Wysocka, T. Swigut, H. Xiao, T.A. Milne, S.Y. Kwon, J. Landry, M. Kauer, A.J. Tackett, B.T. Chait, P. Badenhorst, C. Wu, C.D. Allis, A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling, *Nature* 442 (2006) 86–90, <https://doi.org/10.1038/nature04815>.
- [220] H. Li, S. Ilin, W. Wang, E.M. Duncan, J. Wysocka, C.D. Allis, D.J. Patel, Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF, *Nature* 442 (2006) 91–95, <https://doi.org/10.1038/nature04802>.
- [221] S.G. Alkhatib, J.W. Landry, The nucleosome remodeling factor, *FEBS Lett.* 585 (2011) 3197–3207, <https://doi.org/10.1016/j.febslet.2011.09.003>.
- [222] A.J. Ruthenburg, H. Li, T.A. Milne, S. Dewell, R.K. McGinty, M. Yuen, B. Ueberheide, Y. Dou, T.W. Muir, D.J. Patel, C.D. Allis, Recognition of a mono-nucleosomal histone modification pattern by BPTF via multivalent interactions, *Cell* 145 (2011) 692–706, <https://doi.org/10.1016/j.cell.2011.03.053>.
- [223] J.F. Flanagan, L.Z. Mi, M. Chruszcz, M. Cymborowski, K.L. Clines, Y. Kim, W. Minor, F. Rastinejad, S. Khorasanizadeh, Double chromodomains cooperate to recognize the methylated histone H3 tail, *Nature* 438 (2005) 1181–1185, <https://doi.org/10.1038/nature04290>.
- [224] R.J. Sims, C.F. Chen, H. Santos-Rosa, T. Kouzarides, S.S. Patel, D. Reinberg, Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains, *J. Biol. Chem.* 280 (2005) 41789–41792, <https://doi.org/10.1074/jbc.C500395200>.
- [225] J.J. Lin, L.W. Lehmann, G. Bonora, R. Sridharan, A.A. Vashisht, N. Tran, K. Plath, J.A. Wohlschlegel, M. Carey, Mediator coordinates PIC assembly with recruitment of CHD1, *Genes Dev.* 25 (2011) 2198–2209, <https://doi.org/10.1101/gad.17554711>.
- [226] A. Gaspar-Maia, A. Alajem, F. Polesso, R. Sridharan, M.J. Mason, A. Heidersbach, J. Ramalho-Santos, M.T. McManus, K. Plath, E. Meshorer, M. Ramalho-Santos, Chd1 regulates open chromatin and pluripotency of embryonic stem cells, *Nature* 460 (2009) 863–868, <https://doi.org/10.1038/nature08212>.
- [227] A.B. Patel, R.K. Louder, B.J. Greber, S. Grünberg, J. Luo, J. Fang, Y. Liu, J. Ranish, S. Hahn, E. Nogales, Structure of human TFIID and mechanism of TBP loading onto promoter DNA, *Science* 362 (2018) eaau8872, <https://doi.org/10.1126/science.aau8872>.
- [228] T. Bhuiyan, H.T.M. Timmers, Promoter recognition: putting TFIID on the spot, *Trends Cell Biol.* 29 (2019) 752–763, <https://doi.org/10.1016/j.tcb.2019.06.004>.
- [229] S.M. Lauberth, T. Nakayama, X. Wu, A.L. Ferris, Z. Tang, S.H. Hughes, R.G. Roeder, H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation, *Cell* 152 (2013) 1021–1036, <https://doi.org/10.1016/j.cell.2013.01.052>.
- [230] R. van Nuland, A.W. Schram, F.M.A. van Schaik, P.W.T.C. Jansen, M. Vermeulen, H.T. Marc Timmers, Multivalent engagement of TFIID to nucleosomes, *PLoS One* 8 (2013) e73495, <https://doi.org/10.1371/journal.pone.0073495>.
- [231] Z. Tang, W.-Y. Chen, M. Shimada, U.T.T. Nguyen, J. Kim, X.-J. Sun, T. Sengoku, R.K. McGinty, J.P. Fernandez, T.W. Muir, R.G. Roeder, SET1 and p300 act synergistically, through coupled histone modifications, in transcriptional activation by p53, *Cell* 154 (2013) 297–310, <https://doi.org/10.1016/j.cell.2013.06.027>.
- [232] Y. Dou, T.A. Milne, A.J. Tackett, E.R. Smith, A. Fukuda, J. Wysocka, C.D. Allis, B.T. Chait, J.L. Hess, R.G. Roeder, Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF, *Cell* 121 (2005) 873–885, <https://doi.org/10.1016/j.cell.2005.04.031>.
- [233] T. Margaritis, V. Oreál, N. Brabers, L. Maestroni, A. Vitaliano-Prunier, J.J. Benschop, S. van Hooff, D. van Leenen, C. Dargemont, V. Géli, F.C.P. Holstege, Two distinct repressive mechanisms for histone 3 lysine 4 methylation through promoting 3'-end antisense transcription, *PLoS Genet.* 8 (2012) e1002952, <https://doi.org/10.1371/journal.pgen.1002952>.
- [234] Y.-G. Gae, G. Gangloff, J.-C. Pointud, S. Thuault, L. Carré, C. Romier, S. Muratoglu, M. Brand, L. Tora, J.-L. Couderc, I. Davidson, The TFIID components human TAF II 140 and Drosophila BIP2 (TAF II 155) are novel metazoan homologues of yeast TAF II 47 containing a histone fold and a PHD finger downloaded from, *Mol. Cell. Biol.* 21 (2001) 5109–5121, <https://doi.org/10.1128/MCB.21.15.5109-5121.2001>.
- [235] S.V. Antonova, J. Boeren, H.T.M. Timmers, B. Snel, Epigenetics and transcription regulation during eukaryotic diversification: the saga of TFIID, *Genes Dev.* 33 (2019) 888–902, <https://doi.org/10.1101/gad.300475.117>.
- [236] T.L. Lenstra, J.J. Benschop, T.S. Kim, J.M. Schulze, N.A.C.H. Brabers, T. Margaritis, L.A.L. van de Pasch, S.A.A.C. van Heesch, M.O. Brok, M.J.A. Groot Koerkamp, C.W. Ko, D. van Leenen, K. Sameith, S.R. van Hooff, P. Lijnzaad, P. Kemmeren, T. Hentrich, M.S. Kobor, S. Buratowski, F.C.P. Holstege, The specificity and topology of chromatin interaction pathways in yeast, *Mol. Cell* 42 (2011) 536–549, <https://doi.org/10.1016/j.molcel.2011.03.026>.
- [237] S. Ramakrishnan, S. Pokhrel, S. Palani, C. Pflueger, T.J. Parnell, B.R. Cairns, S. Bhaskara, M.B. Chandrasekharan, Counteracting H3K4 methylation modulators Set1 and Jhd2 co-regulate chromatin dynamics and gene transcription, *Nat. Commun.* 7 (2016) 11949, <https://doi.org/10.1038/ncomms11949>.
- [238] J.K. Choi, Contrasting chromatin organization of CpG islands and exons in the human genome, *Genome Biol.* 11 (2010) R70, <https://doi.org/10.1186/gb-2010-11-7-r70>.
- [239] A. Weiner, H.V. Chen, C.L. Liu, A. Rahat, A. Klien, L. Soares, M. Gudipati, J. Pfeiffer, A. Regev, S. Buratowski, J.A. Pleiss, N. Friedman, O.J. Rando, Systematic dissection of roles for chromatin regulators in a yeast stress response, *PLoS Biol.* 10 (2012) e1001369, <https://doi.org/10.1371/journal.pbio.1001369>.
- [240] S. Boa, C. Coert, H.G. Patterson, *Saccharomyces cerevisiae* Set1 p is a methyltransferase specific for lysine 4 of histone H3 and is required for efficient gene

- expression, *Yeast* 20 (2003) 827–835, <https://doi.org/10.1002/yea.995>.
- [241] C. Nislow, E. Ray, L. Pillus, SET1, a yeast member of the Trithorax family, functions in transcriptional silencing and diverse cellular processes, *Mol. Biol. Cell* 8 (1997) 2421–2436, <https://doi.org/10.1091/mbc.8.12.2421>.
- [242] G.M. Martin, D.A. King, E.M. Green, P.E. Garcia-Nieto, R. Alexander, S.R. Collins, N.J. Krogan, O.P. Gozani, A.J. Morrison, Set5 and Set1 cooperate to repress gene expression at telomeres and retrotransposons, *Epigenetics* 9 (2014) 513–522, <https://doi.org/10.4161/epi.27645>.
- [243] T. Miller, N.J. Krogan, J. Dover, H. Erdjument-Bromage, P. Tempst, M. Johnston, J.F. Greenblatt, A. Shilatifard, COMPASS: a complex of proteins associated with a trithorax-related SET domain protein, *Proc. Natl. Acad. Sci. U. S. A.* 98 (2001) 12902–12907, <https://doi.org/10.1073/pnas.231473398>.
- [244] L. Wang, C.K. Collings, Z. Zhao, K.A. Cozzolino, Q. Ma, K. Liang, S.A. Marshall, C.C. Sze, R. Hashizume, J.N. Savas, A. Shilatifard, A cytoplasmic COMPASS is necessary for cell survival and triple-negative breast cancer pathogenesis by regulating metabolism, *Genes Dev.* 31 (2017) 2056–2066, <https://doi.org/10.1101/gad.306092.117>.
- [245] T. Clouaire, S. Webb, P. Skene, R. Illingworth, A. Kerr, R. Andrews, J.H. Lee, D. Skalnik, A. Bird, Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells, *Genes Dev.* 26 (2012) 1714–1728, <https://doi.org/10.1101/gad.194209.112>.
- [246] T.M. Franks, A. McCloskey, M. Shokirev, C. Benner, A. Rathore, M.W. Hetzer, Nup98 recruits the Wdr82-Set1A/COMPASS complex to promoters to regulate H3K4 trimethylation in hematopoietic progenitor cells, *Genes Dev.* 31 (2017) 2222–2234, <https://doi.org/10.1101/gad.306753.117>.
- [247] D. Cano-Rodriguez, R.A.F. Gjaltema, L.J. Jilderda, P. Jellema, J. Dokter-Fokkens, M.H.J. Ruijters, M.G. Rots, Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner, *Nat. Commun.* 7 (2016) 12284, <https://doi.org/10.1038/ncomms12284>.
- [248] H. Jiang, A. Shukla, X. Wang, W. Chen, B.E. Bernstein, R.G. Roeder, Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains, *Cell* 144 (2011) 513–525, <https://doi.org/10.1016/j.cell.2011.01.020>.
- [249] K. Cao, C.K. Collings, S.A. Marshall, M.A. Morgan, E.J. Rendleman, L. Wang, C.C. Sze, T. Sun, E.T. Bartom, A. Shilatifard, SET1A/COMPASS and shadow enhancers in the regulation of homeotic gene expression, *Genes Dev.* 31 (2017) 787–801, <https://doi.org/10.1101/gad.294744.116>.
- [250] B.A. Benayoun, E.A. Pollina, D. Ucar, S. Mahmoudi, K. Karra, E.D. Wong, K. Devarajan, A.C. Daugherty, A.B. Kundaje, E. Mancini, B.C. Hitz, R. Gupta, T.A. Rando, J.C. Baker, M.P. Snyder, J.M. Cherry, A. Brunet, H3K4me3 breadth is linked to cell identity and transcriptional consistency, *Cell* 158 (2014) 673–688, <https://doi.org/10.1016/j.cell.2014.06.027>.
- [251] B.D. Yu, J.L. Hess, S.E. Horning, G.A.J. Brown, S.J. Korsmeyer, Altered Hox expression and segmental identity in Mll-mutant mice, *Nature* 378 (1995) 505–508, <https://doi.org/10.1038/378505a0>.
- [252] H. Yagi, K. Deguchi, A. Aono, Y. Tani, T. Kishimoto, T. Komori, Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice, *Blood* 92 (1998) 108–117, [https://doi.org/10.1182/blood.v92.1.108.413k11\\_108.117](https://doi.org/10.1182/blood.v92.1.108.413k11_108.117).
- [253] P. Ernst, J.K. Fisher, W. Avery, S. Wade, D. Foy, S.J. Korsmeyer, Definitive hematopoiesis requires the mixed-lineage leukemia gene, *Dev. Cell* 6 (2004) 437–443, [https://doi.org/10.1016/S1534-5807\(04\)00061-9](https://doi.org/10.1016/S1534-5807(04)00061-9).
- [254] S. Glaser, J. Schaft, S. Lubitz, K. Vintersten, F. van der Hoeven, K.R. Tufteland, R. Aasland, K. Anastasiadis, S.-L. Ang, A.F. Stewart, Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development, *Development* 133 (2006) 1423–1432, <https://doi.org/10.1242/dev.02302>.
- [255] R.J. Klose, S. Cooper, A.M. Farcas, N.P. Blackledge, N. Brockdorff, Chromatin sampling—an emerging perspective on targeting polycomb repressor proteins, *PLoS Genet.* 9 (2013) e1003717, <https://doi.org/10.1371/journal.pgen.1003717>.
- [256] I.B. Dodd, M.A. Micheelsen, K. Sneppen, G. Thon, Theoretical analysis of epigenetic cell memory by nucleosome modification, *Cell* 129 (2007) 813–822, <https://doi.org/10.1016/j.cell.2007.02.053>.
- [257] A. Angel, J. Song, C. Dean, M. Howard, A Polycomb-based switch underlying quantitative epigenetic memory, *Nature* 476 (2011) 105–109, <https://doi.org/10.1038/nature10241>.
- [258] P.A. Steffen, L. Ringrose, What are memories made of? How polycomb and trithorax proteins mediate epigenetic memory, *Nat. Rev. Mol. Cell Biol.* 15 (2014) 340–356, <https://doi.org/10.1038/nrm3789>.
- [259] S. Berry, C. Dean, M. Howard, Slow chromatin dynamics allow polycomb target genes to filter fluctuations in transcription factor activity, *Cell Syst* 4 (2017) 445–457.e8, <https://doi.org/10.1016/j.cels.2017.02.013>.
- [260] H. Yang, M. Howard, C. Dean, Antagonistic roles for H3K36me3 and H3K27me3 in the cold-induced epigenetic switch at Arabidopsis FLC, *Curr. Biol.* 24 (2014) 1793–1797, <https://doi.org/10.1016/j.cub.2014.06.047>.
- [261] K. Sneppen, I.B. Dodd, A simple histone code opens many paths to epigenetics, *PLoS Comput. Biol.* 8 (2012) e1002643, <https://doi.org/10.1371/journal.pcbi.1002643>.
- [262] H. Binder, L. Steiner, J. Przybilla, T. Rohlf, S. Prohaska, J. Galle, Transcriptional regulation by histone modifications: towards a theory of chromatin re-organization during stem cell differentiation, *Phys. Biol.* 10 (2013) 026006, <https://doi.org/10.1088/1478-3975/10/2/026006>.
- [263] T. Rohlf, L. Steiner, J. Przybilla, S. Prohaska, H. Binder, J. Galle, Modeling the dynamic epigenome: from histone modifications towards self-organizing chromatin, *Epigenomics* 4 (2012) 205–219, <https://doi.org/10.2217/epi.11.117>.
- [264] T.L. Lenstra, J. Rodriguez, H. Chen, D.R. Larson, Transcription dynamics in living cells, *Annu. Rev. Biophys.* 45 (2016) 25–47, <https://doi.org/10.1146/annurev-biophys-062215-010838>.
- [265] K. Sneppen, L. Ringrose, Theoretical analysis of Polycomb-Trithorax systems predicts that poised chromatin is bistable and not bivalent, *Nat. Commun.* 10 (2019) 2133, <https://doi.org/10.1038/s41467-019-10130-2>.
- [266] F.S. Howe, H. Fischl, S.C. Murray, J. Mellor, Is H3K4me3 instructive for transcription activation? *BioEssays* 39 (2017) 1–12, <https://doi.org/10.1002/bies.201600095>.
- [267] A.H. Turberfield, T. Kondo, M. Nakayama, Y. Koseki, H.W. King, H. Koseki, R.J. Klose, KDM2 proteins constrain transcription from CpG island gene promoters independently of their histone demethylase activity, *Nucleic Acids Res.* 47 (2019) 9005–9023, <https://doi.org/10.1093/nar/gkz607>.