

Prospects & Overviews

Is H3K4me3 instructive for transcription activation?

Françoise S. Howe, Harry Fischl, Struan C. Murray and Jane Mellor*

Tri-methylation of lysine 4 on histone H3 (H3K4me3) is a near-universal chromatin modification at the transcription start site of active genes in eukaryotes from yeast to man and its levels reflect the amount of transcription. Because of this association, H3K4me3 is often described as an activating histone modification and assumed to have an instructive role in the transcription of genes, but the field is lacking a conserved mechanism to support this view. The overwhelming finding from genome-wide studies is that actually very little transcription changes upon removal of most H3K4me3 under steady-state or dynamically changing conditions, including at mammalian CpG island promoters. Instead, rather than a major role in instructing transcription, time-resolved experiments provide more evidence supporting the deposition of H3K4me3 into chromatin as a result of transcription, influencing processes such as memory of previous states, transcriptional consistency between cells in a population and transcription termination.

Keywords:

■ CFP1/Spp1; chromatin; H3K4me3; methylation; Set1; transcription; transcription activation

Introduction

The most highly cited paper, with >1,000 citations, under the search term 'K4' or 'Set1' is that from the Kouzarides group [1],

with the title 'Active genes are tri-methylated at K4 of histone H3'. This was one of the first papers to propose a link between the level of H3K4me3 on a gene and the amount of transcript produced. Indeed, H3K4me3 is commonly referred to as an 'activating' histone modification. But whilst there is clearly a strong association between H3K4me3, active genes and RNA levels, for nearly 15 years many groups have searched for causality in this relationship and failed to find any overarching mechanism for how H3K4me3 is instructive for transcription activation. Here we aim to re-evaluate the evidence for whether H3K4me3 actually regulates transcription initiation and whether the failure to find an overarching mechanism reflects distinct functions in different organisms.

H3K4 is methylated by Set1 or Set1-like methyltransferase complexes

In eukaryotes, DNA is packaged into chromatin, with 147 bp DNA wrapped around a core of eight histone proteins to form a nucleosome [2]. The tails of these histone proteins protrude from the nucleosome core and are extensively post-translationally modified. Histone H3 lysine 4 (H3K4) can be modified by the addition of one, two, or three methyl groups and it is likely that the levels and distribution of H3K4 methylation will reflect the combined action of methyltransferases and demethylases [3–5]. The enzymes that methylate H3K4 are highly conserved between yeast, worms, plants, flies, and humans [6]. They each contain a SET methyltransferase domain and are present in a complex with accessory proteins that are both required for lysine methyltransferase (KMT) activity and for influencing the state (mono-, di-, or tri-) of methylation deposited. In the yeast *Saccharomyces cerevisiae*, H3K4 is methylated solely by the lysine methyltransferase Set1/KMT2 in a complex containing seven accessory subunits: Swd1, Swd3, Swd2, Sdc1, Bre2, Spp1 and Shg1 (Fig. 1) [6]. *Drosophila melanogaster* has three Set1 homologues (dSet1, *Trithorax* and *Trithorax-related*) whereas mammals have six complexes that catalyse H3K4 methylation (SET1A/B and MLL1-4) (Fig. 1) [6]. These complexes are not redundant, as demonstrated by the embryonic lethality caused by deletion of individual MLL genes [7–9]. The *Drosophila* dSet1 complex and the two mammalian SET1

DOI 10.1002/bies.201600095

Department of Biochemistry, University of Oxford, Oxford, UK

*Corresponding author:

Jane Mellor
E-mail: jane.mellor@bioch.ox.ac.uk

Abbreviations:

(m)ESC, (murine) embryonic stem cell; H3K4me3, tri-methylation of lysine 4 on histone H3; NDR, nucleosome-depleted region; TSS, transcription start site.

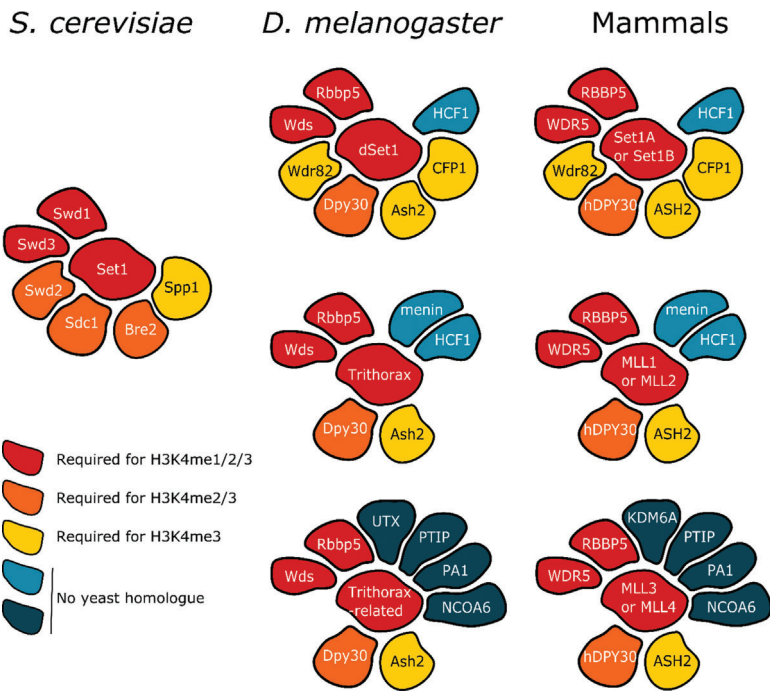


Figure 1. The Set1-related complexes in yeast, *Drosophila* and humans. There is one Set1 complex in yeast but three in *Drosophila* and six in mammals. Homologous subunits are the same shapes and similarly positioned relative to the central methyltransferase protein. Subunits that are required for H3K4me3, H3K4me2/3 or all states of H3K4 methylation are shown in yellow, orange and red, respectively. *D. melanogaster* and mammals have complex-specific subunits and these are shown in blue.

complexes are most similar to the yeast complex (Fig. 1) [6] and are responsible for the bulk of H3K4me3 [10–12]. The *Trithorax*, *Trithorax-related* and MLL complexes are Set1 complex-like and consist of similar core components (Fig. 1), although the structural and functional organisation of these subunits differs from the yeast complex [13, 14]. Additionally, these Set1-like complexes appear to have more specific functions: MLL1 and MLL2 methylate H3K4 at different subsets of homeotic genes [8, 15] whereas MLL3/4 are linked to H3K4me1 and enhancer function [16]. Set1-related complexes have been implicated in cancer in mammals [17], ageing in *C. elegans* [18], embryonic development [19, 20] and stem cell biology [21–23].

H3K4me3 peaks at transcription start sites

Using chromatin immunoprecipitation (ChIP), the distribution of H3K4 methylation has been mapped in a range of organisms [1, 4, 24–35]. Whilst H3K4 methylation is predominantly associated with genes, leading to a strong correlation with transcription (Table 1), H3K4me3 also functions in DNA repair [36, 37] and targeting of double strand breaks during meiotic recombination [38–40]. Despite the differences between yeast and metazoan gene length and structure, the most striking feature of the H3K4me3 gene distribution, conserved in yeast, plants, worms, flies and mammals, is the peak around the transcription start sites (TSSs) of genes [1, 4, 24–35] (Fig. 2A). Moreover, closely spaced divergent

Table 1. Correlations between H3K4 methylation and gene expression over gene bodies in *S. cerevisiae*

H3K4 methylation state (H3-normalised)	Correlation with	Spearman correlation	References
me1	Transcript levels	–0.318	[29, 70, 127]
	Transcription (NET-seq)	–0.414	
	Transcription (PRO-seq)	–0.163	
me2	Transcript levels	+0.270	[29, 70, 127]
	Transcription (NET-seq)	+0.351	
	Transcription (PRO-seq)	+0.154	
me3	Transcript levels	+0.603	[33, 70, 127]
	Transcription (NET-seq)	+0.738	
	Transcription (PRO-seq)	+0.381	
Strain	Number of genes with changes	p-value	
Deletion of <i>SET1</i> , core Set1C components or H3K4R point mutation ^a	69 increase (transcripts)	<0.01; >1.7-fold change	[43]
	20 decrease (transcripts)	<0.01; >1.7-fold change	[43]
<i>spp1</i> Δ (results in loss of H3K4me3)	57 increase (NET-seq)	<0.05	Harry Fischl, unpublished
	23 decrease (NET-seq)	<0.05	Harry Fischl, unpublished

^aSet1 is epistatic to H3K4R showing all Set1 effects on RNA levels is mediated via H3K4.

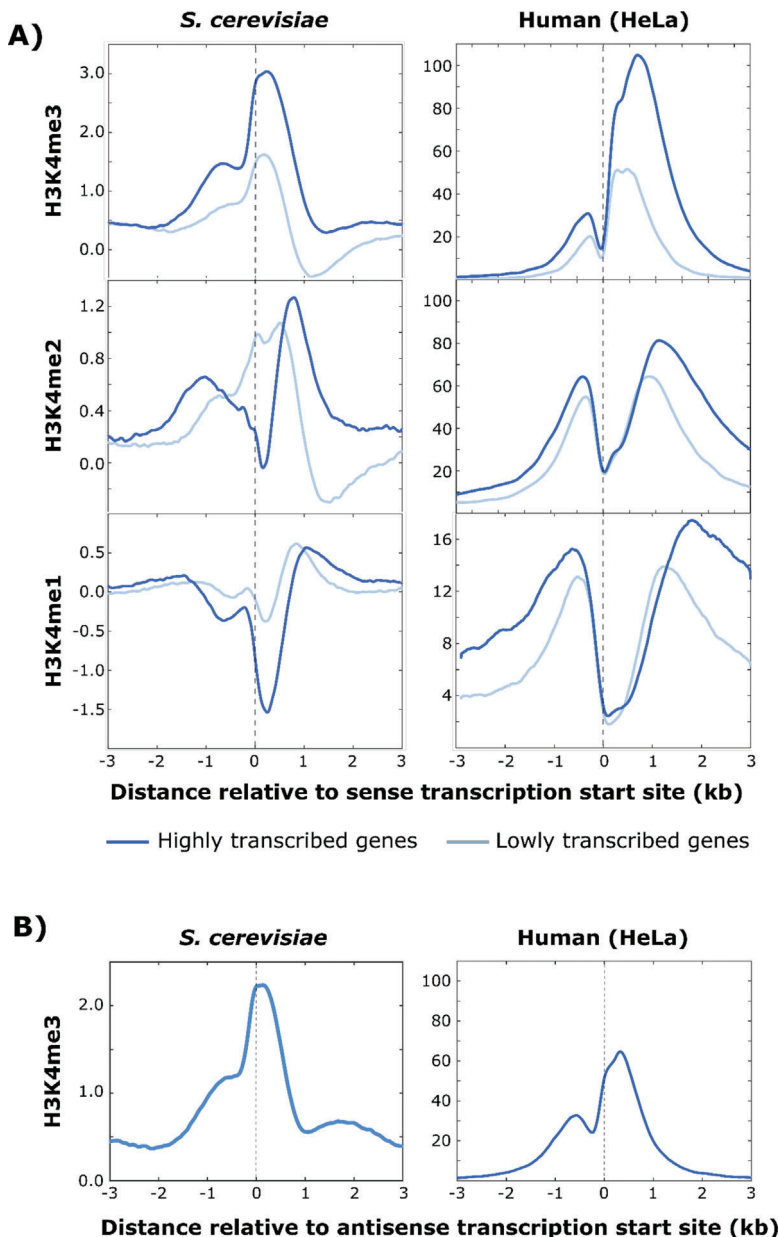


Figure 2. Distribution of H3K4 methylation on genes. **A:** The distribution of H3K4 methylation on yeast genes (left) and mammalian (HeLa) genes (right) 3000 bp upstream or downstream of the sense transcription start site (TSS, dashed line) for highly expressed genes (dark blue) or poorly expressed genes (light blue). For yeast, the top ($n = 1155$) or bottom ($n = 1163$) quintiles of genes in terms of NET-seq reads [70] are used to map the distribution of histone H3-normalised H3K4me1, 2 and 3 [29]. For HeLa cells, only genes with a strong 5' CAGE signal and thus a well-defined TSS were used for the analysis, and the top ($n = 1864$) and bottom ($n = 1864$) quintiles of these genes in terms of mNET-seq reads [139] are used to map the distribution of H3K4me1, 2 and 3 [34]. **B:** As (A) except H3K4me3 distribution on all genes with a defined antisense TSS, aligned to the antisense TSS (yeast $n = 1529$, HeLa $n = 2468$).

promoters, whether for coding or non-coding transcription, have an additional upstream peak of H3K4me3 separated by a nucleosome-depleted region (NDR) [41, 42]. Interestingly, the promoters of antisense transcription at the 3' ends of yeast and

human genes are also marked with a peak of H3K4me3, indicating that this modification is not specific to sense TSSs (Fig. 2B) [43, 44].

As expected, the amount of H3K4me3 on both yeast and mammalian genes is related to the levels of nascent sense transcription, with highly expressed genes having more H3K4me3 (Fig. 2, Table 1). Unlike initially reported [4], both H3K4me3 peak height and peak breadth correlate positively with expression in a range of organisms [35, 45]. Broad H3K4me3 peaks are also an epigenetic signature of tumour suppressor genes and are associated with increased transcription elongation and enhancer activity [35]. Furthermore, tumour cells have narrower H3K4me3 peaks and a lower level of tumour suppressor transcripts. The distributions of H3K4me2 and H3K4me1 are distinct from H3K4me3 (Fig. 2A) and the differing correlations with transcription (Table 1) suggest that there are methylation state-specific functions.

The difficulty of studying the functions associated with different H3K4 methylation states

The study of histone modifications in vivo is complicated by the fact that it is difficult to genetically affect one modification in isolation without being at risk of indirect effects caused by changing the levels of other marks. For example, deletion of yeast *SET1* removes all three states of H3K4 methylation, giving potentially conflated results if each state has different functions. Further conflation comes from the fact that Set1, like many histone methyltransferases, has non-histone substrates, such as the kinetochore protein, Dam1 [46]. Similarly, in *Arabidopsis thaliana*, the H3K4 methyltransferase ATX1 has effects on transcription that are independent of its catalytic activity towards H3K4 [47]. A commonly used approach in yeast and *Dictyostelium discoideum* is to substitute H3K4 with a non-methylatable residue such as alanine [48, 49].

However, this again is not methylation state-specific and also removes H3K4 acetylation [33], which would result in effects that could not be attributed solely to H3K4 methylation. Furthermore, this approach is difficult in other organisms due to the high repeat number for histone genes. An alternative strategy to specifically alter H3K4me3 levels is to interfere with the integrity of the methyltransferase complexes.

Spp1 and CFP1 are required for the majority of H3K4me3 in yeast and mammals

Both yeast Set1 and human SET1A/B complexes have subunits, Spp1 and CFP1 respectively, that when ablated specifically reduce H3K4me3 with little effect on global levels of H3K4me1 or me2 [50–52]. In yeast, reduced H3K4me3, by deletion of *SPP1*, is an important trigger for cell death

(apoptosis) and is linked to poor chronological ageing [53] while absence of CFP1 causes peri-implantation embryonic lethality in mice [54], although murine embryonic stem cells (mESCs) are viable but unable to differentiate [55].

CFP1 is recruited to the ~70% of mammalian promoters that contain a CpG island [52, 56], regions of high CpG density that are not subject to DNA methylation, via a Zn finger CXXC domain that binds to unmethylated CpG dinucleotides [57]. CFP1 can also be recruited to chromatin via its PHD domain that can bind all three H3K4 methylation states [58]. CFP1 is not in the MLL complexes, but its binding functions may be recapitulated by the CXXC domains in MLL1/2 [59, 60] and the third PHD domain in MLL1 [61]. Both SET1A/B and MLL2 complexes act on CpG island promoters, with SET1A/B methylating H3K4 on more highly expressed genes [62] whereas MLL2 preferentially methylates H3K4 at more lowly expressed bivalent gene promoters in embryonic stem cells (ESCs) [63, 64].

Yeast Spp1, like CFP1, has a PHD finger that recognises H3K4me2 and H3K4me3 [65], suggesting that Spp1, through its PHD domain, may bind H3K4me2 and stabilise the Set1 complex on chromatin to bring about H3K4me3 [29]. Another possibility is that, by binding H3K4me2/3 on one histone tail, Spp1 can recruit Set1 to methylate the other histone H3 tail in the same nucleosome or neighbouring nucleosomes. It is common for components of histone modifying enzyme complexes to play a role in the deposition of a modification while also containing a domain that binds to that modification. Additionally, it was speculated that the binding of Spp1 to H3K4me2/3 may protect these modifications from demethylation [29].

Although Spp1 directly interacts with the Set1 N-SET catalytic domain [66, 67], in vitro studies demonstrate that Spp1 is not required for tri-methylation of H3K4 on free histone H3, arguing against a strict enzymatic requirement of Spp1 for Set1-mediated H3K4me3 [66, 67]. In agreement, in vivo, ~20% global H3K4me3 remains after *SPP1* deletion [3, 40]. Since this residual amount is unevenly distributed between individual genes, with H3K4me3 on some genes such as the ribosomal protein genes being largely independent of Spp1 [3], it is possible to restrict analysis to the class of genes with H3K4me3, the most affected by *SPP1* deletion. Similarly, whilst CFP1 only recruits SET1A/B complexes to CpG island promoters, analysis could be limited to these genes when ascertaining any effect of H3K4me3.

Loss of H3K4me3 has minor effects on global transcription

Deletion of CFP1 from mESCs causes loss of most H3K4me3 from expressed CpG island-associated genes but yet there are minimal changes in transcript levels at these genes, assessed by microarray, or nascent transcription as measured by levels of RNA polymerase II by ChIP and genomic run-on sequencing [62]. Furthermore, only a minority of the DNA damage-inducible genes that gain H3K4me3 in a CFP1-dependent manner showed defects in transcript levels in *cfp1*^{-/-} mESCs [68]. However, it is plausible that the basal level of H3K4me3 remaining in *cfp1*^{-/-} cells is sufficient for

transcription activation. Interestingly, small defects in repression upon DNA damage were also observed at some genes in *cfp1*^{-/-} mESCs, associating H3K4me3 with both gene-specific activation and repression. Additionally, levels of both divergent upstream transcripts and antisense transcripts arising from the 3' regions of the DNA damage-inducible genes are significantly increased in the absence of CFP1 [68]. These data support a role for CFP1, and possibly H3K4me3, in repressing divergent and convergent antisense transcripts at the 755 genes induced by DNA damage.

Similar to mammalian systems, deletion of *SPP1* from yeast, resulting in gene-specific H3K4me3 reduction [3], has virtually no effect on steady-state or dynamically changing mRNAs [43, 69] or transcription (Table 1 and Fig. 3A, unpublished) assessed using native elongating transcript analysis, NET-seq [70]. We find only 57 (1.0%) genes are significantly ($p < 0.05$) upregulated and 23 (0.4%) genes are significantly downregulated at the level of transcription upon deletion of *SPP1* at steady-state growth. To ask whether the genes that lose the most H3K4me3 are those whose transcription changes the most, we made use of a ChIP-seq dataset measuring H3K4me3 in a strain in which levels of H3K4me3 are affected similarly to those in the *spp1Δ* strain [3]. We found no correlation ($r_s = 0.006$) between the change in H3K4me3 and the change in transcription upon deletion of *SPP1* at the +1 nucleosome where H3K4me3 is most abundant. Moreover, when we studied a dynamic system during a change from glucose-containing to galactose-containing growth medium, a carbon source shift shown to alter the transcription of ~20% of yeast genes [71], in the *spp1Δ* strain fewer than 15 genes are significantly ($p < 0.05$) differentially transcribed (unpublished data). Of note, the induction profiles of the *GAL* genes that are strongly upregulated upon shift to galactose-containing medium, are not altered in the *spp1Δ* strain. Yeast H3K4me3 levels can also be controlled by the JmJc family demethylase Jhd2 [72–75], which is capable of removing all three H3K4 methylation states in vivo [76]. Analysis of datasets from the Holstege [69] and Yan labs (accession number GSE67211; unpublished) shows that genes with increased H3K4me3 upon deletion of *JHD2* do not have higher levels of mRNA (Fig. 3B and C).

Context-dependent effects of H3K4me3 on gene expression

The results from the deletion of CFP1 in mESCs and *SPP1* in yeast indicate that, whilst at the majority of genes no effect on transcription is seen, H3K4me3 can be associated with both activation and repression of transcription at a small subset of genes. For example, ectopic deposition of H3K4me3, either by mis-regulation of the SET1A/B complex in *cfp1*^{-/-} mESCs or by targeting of a meiotic H3K4 methyltransferase PRDM9 to sites normally depleted for H3K4me3 in human cells can lead to transcription of these regions [62, 77], whereas H3K4me3 increases at yeast ribosomal protein genes during their repression in response to diamide stress [78, 79]. This apparent disparity between activating and repressing functions can be resolved by the fact that H3K4me3 can recruit a number of factors, many of which have been linked to

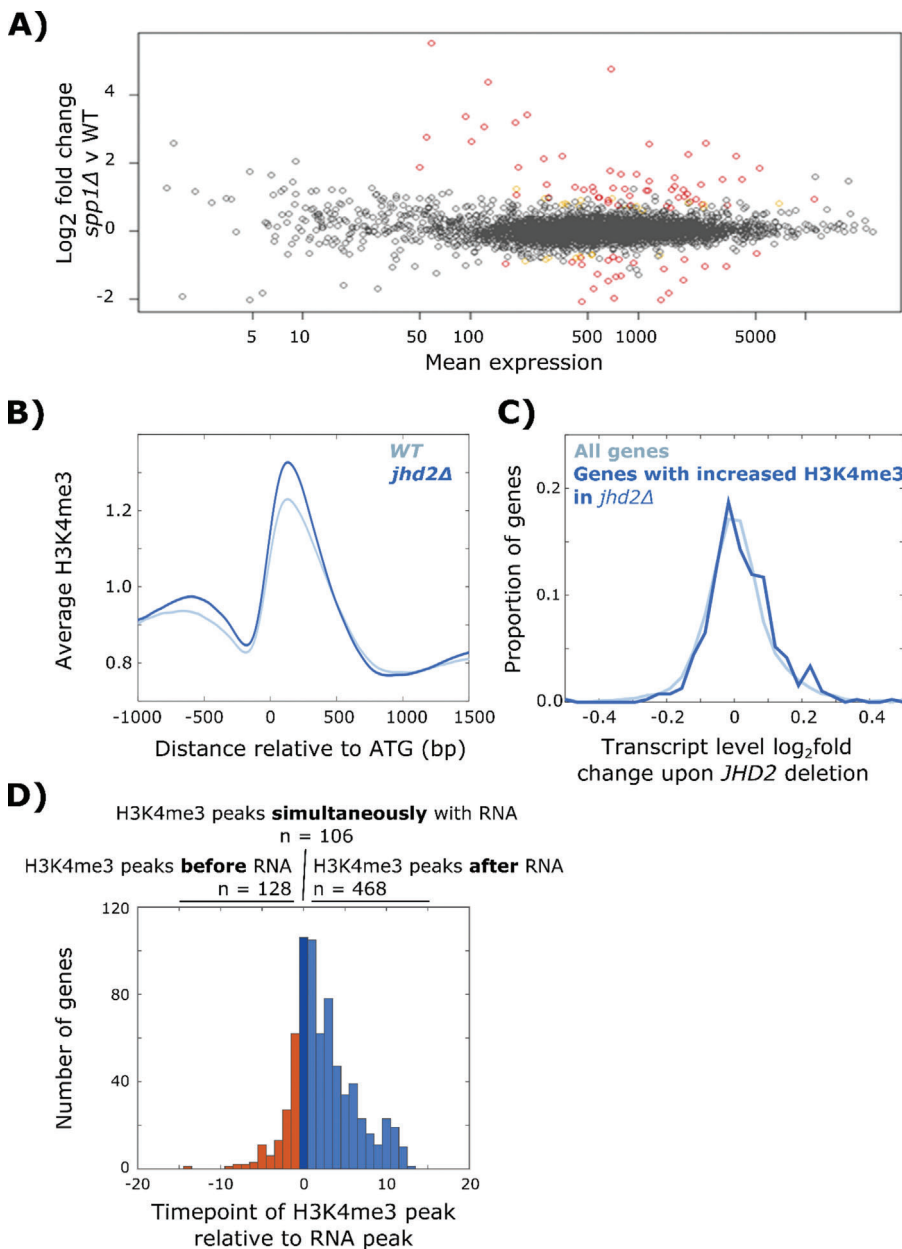


Figure 3. *SPP1* and *JHD2* have minimal effects on transcription or transcript levels in *S. cerevisiae*. **A:** Log₂ fold change in transcription (as measured by NET-seq) in the *spp1Δ* versus the wild type strain at yeast transcription units ($n = 5579$) plotted against the mean expression level for each gene. Genes showing significant changes are shown in red ($p < 0.01$) or yellow ($p < 0.05$). **B:** The average H3K4me3 ChIP profile for 5774 yeast genes in the wild type (light blue) and *jhd2Δ* (dark blue) strains (Yan lab, GSE67211). **C:** Despite the increase in H3K4me3 at 393 genes in the *jhd2Δ* strain (B), there are no significant changes in transcript levels of these genes (dark blue) compared to the change at all genes (light blue, $n = 5774$) [69]. **D:** Histogram showing the timing of the H3K4me3 peak relative to the transcript peak for each gene whose transcript peaks during oxidative phase of the yeast metabolic cycle ($n = 702$) [100]. Time labels refer to the difference between the 16 sampling time-points used by Kuang et al. Positive time values (light blue) indicate genes at which H3K4me3 peaks after the RNA.

transcription (Table 2), via protein folds such as the PHD finger, chromo-domain or Tudor domain [80]. The factor recruited may depend on the presence or absence of neighbouring modifications for example, BPTF in the NURF chromatin remodelling complex has both a PHD finger and a bromo-domain and preferentially binds H3K4me3 when H4K16 on the same nucleosome is acetylated [81]. Conversely, Spp1 binding to H3K4me3 is inhibited by asymmetric di-methylation of neighbouring H3R2 [29]. Thus depending on which factor and other co-factors are recruited, the residency time and the affinity of those interactions, H3K4me3 has the potential to activate or repress transcription at different subsets of genes, illustrated by the capacity of H3K4me3 to recruit both histone acetyltransferases and deacetylases (Table 2). However, it should be noted that the conclusions of the studies that link these recruited factors to transcriptional activation or repression often rely on the consequential effect on other histone modifications that also correlate with transcription, for example, histone acetylation, without definitively demonstrating the requirement of the factor interaction with H3K4me3 for this transcriptional activation/repression. Furthermore, a number of the so-called 'activating' factors that specifically bind to H3K4me3 have subsequently been shown to also have repressive functions, for example, the chromatin remodeller yeast Isw1/human SNF2h [82–84].

A robust study in human colorectal carcinoma HCT116 cells has demonstrated a mechanism by which

H3K4me3 promotes transcription activation at the subset of p53-dependent DNA damage-inducible genes whose transcription is affected by H3K4me3 loss [85]. This involves the TFIID subunit TAF3, which binds directly to H3K4me3 [86]. Using high quality biochemical analysis with engineered nucleosomes containing H3K4me3, TAF3 recruitment to H3K4me3 is shown to activate p53-dependent transcription in vitro and in vivo [85]. Additional support for the link between H3K4me3 and TFIID/TAF3 comes from the observation that TBP levels are reduced at the *Il20ra* gene promoter that loses most of its H3K4me3 in *mll1*^{-/-} mouse embryonic fibroblasts (MEFs) [15]. However, this mechanism may only be functional at a small subset of genes: depletion of TAF3 only altered the transcripts from 5.6% of genes [85]. Interestingly, this function of H3K4me3 may be specific to mammalian cells as the yeast TAF3 does not contain a PHD finger [87] and plant

Table 2. Proteins that interact with H3K4me3

Factor	H3K4 methylation state	Interaction domain	Factor is part of complex	Associated consequence of interaction	References
CFP1	me3	PHD	SET1A/B	Increased H3K9ac and H3K14ac upon DNA damage	[58, 68]
Spp1	me2/3	PHD	Set1 complex	Spreading of H3K4me3; initiation of meiotic recombination	[29, 39, 40]
TAF3	me3	PHD	TFIID	Transcriptional activation at p53-regulated genes	[85, 86]
lsw1/ SNF2h	me2/3	Via CHD1 for SNF2h	lsw1a, lsw1b, SNF2h	Repression of transcripts; chromatin remodelling	[82–84]
Yng1	me3	PHD	NuA3	H3K14 acetylation	[114, 115]
Sgf29	me2/3	Tandem Tudor domains	SAGA	H3K9,14,18 acetylation; correlates with gene activation	[26, 128, 129]
PHF2	me2/3	PHD		H3K9 demethylation	[130]
PHF8	me2/3	PHD		H3K9me1/2 demethylation	[131, 132]
PHF13	me2/3	PHD		Recruitment of transcriptional regulators	[133]
Pho23/ Cti6	me3	PHD	Rpd3L	Lysine deacetylation	[113]
Yng2/ ING2	me2/3	PHD	NuA4/Sin3 HDAC	Lysine deacetylation after DNA damage; silences proliferation genes	[134]
CHD1	me3	Chromo		ATP-dependent chromatin remodelling; recruitment of U2 snRNP components	[83, 105, 135]
BPTF	me3	PHD	NURF	ISWI-containing ATP-dependent chromatin remodelling complex; required for HOX gene expression/repression	[136, 137]
Tip60	me3	?	dTip60 complex	Acetylation and exchange of variant H2A; release of paused polymerase at <i>hsp70</i> in <i>D. melanogaster</i>	[138]

TFIID does not have an equivalent TAF3 subunit [88]. Whilst it is very conceivable that H3K4me3 may have different modes of action at distinct subsets of genes and in different cell types and organisms, the positive genome-wide correlation between H3K4me3 and transcription exists in all organisms studied and so there must be other mechanisms to explain this widespread association.

No instructive role for H3K4me3 in transcription?

The fact that at the majority of genes transcription is unaffected by loss of most H3K4me3 argues against a globally instructive role for H3K4me3 in transcription. This is supported by the observation that *in vitro*, H3K4me3 is not required for transcription of a chromatinised substrate [89]. Instead, studies suggest the Set1/Set1-like complexes may rely on pre-existing transcription-dependent events to deposit H3K4me3 into chromatin, ruling out a strictly instructive role but explaining the strong correlation with transcription. Transcription-dependent recruitment of the Set1/MLL proteins has been described for yeast, *D. melanogaster*, *A. thaliana* and mammals [90–94]. In yeast, the Paf1 elongation factor is required for the recruitment of the Set1 complex to RNA polymerase II post-initiation [90], although evidence disputes this requirement in mammals [93, 95]. Additionally, in yeast, co-transcriptionally deposited histone H2B ubiquitination [96] is required for H3K4me3 [97, 98], which acts in part by preventing Jhd2-dependent demethylation [99]. Further evidence for the deposition of H3K4me3 as a result of

transcription comes from time-resolved experiments. The study of the oscillating expression of over half of all yeast genes in the yeast metabolic cycle reveals that at the majority of these genes, H3K4me3 peaks after the peak of the RNA transcript, especially for those genes whose transcripts peak during the oxidative phase (Fig. 3D) [100]. A delay of H3K4me3 relative to induced RNA is also observed during meiosis in yeast [38] and removal of H3K4 methylation does not have a strong effect on meiotic gene induction [101]. Similarly, for genes varying expression during diurnal cycles in the mouse liver, H3K4me3 on average peaked 1.3 h after RNA polymerase II on chromatin [102] and bulk H3K4me3 is not observed before the start of transcription during the zebrafish maternal-zygotic transition [103].

Another possibility is that, although H3K4me3 and transcription correlate strongly, these are actually independent events. A simple explanation is that an NDR over DNA with a suitable base composition, such as a CpG island in mammals, recruits SET1A/B leading to H3K4me3. Independently, an NDR might also be sufficient to recruit RNA polymerase II, either supported by DNA-bound transcription factors or not, leading to transcription, hence the apparent link between H3K4me3 and transcription. However, as the non-perfect correlation between H3K4me3 and transcription suggests, it is possible to get H3K4me3 without transcription and vice versa. Evidence for this comes from mESCs: at engineered CpG islands with no nearby promoters, CFP1-dependent H3K4me3 can be deposited when surrounding transcription levels, assessed using RNA polymerase II ChIP, are very low or absent and this H3K4me3 does not, in turn,

lead to transcription [52]. Furthermore, in *Xenopus* embryos, developmentally regulated hypomethylated but not hypermethylated CpG island promoters have H3K4me3 hours before their transcriptional activation [104]. In these contexts, H3K4me3 deposition appears to be solely a function of the unmethylated CpG content of the DNA, allowing specific recruitment of CFP1 and SET1A/B directly.

The evidence presented so far has suggested that whilst H3K4me3 can influence the transcription of some genes positively and negatively, H3K4me3 can also be deposited after or independently of transcription. In this situation, the main function of H3K4me3 is unlikely to be regulation of transcription initiation, supported by the fact that the absence of the majority of H3K4me3 does not have pervasive effects on nascent transcription. So if H3K4me3 is not affecting transcription globally, what other functions might it have?

A role for H3K4me3 in splicing pre-mRNAs?

One of the H3K4me3-binding proteins in humans is CHD1 [105] (Table 2). CHD1 is linked to splicing as it is able to stably associate with U2 snRNP in the spliceosome and depletion of CHD1 decreases splicing efficiencies eightfold at early time-points of *IRF1* gene induction in vivo [83]. However, the direct link between CHD1 association with H3K4me3 and splicing is challenged, as the splicing defect is only threefold upon ASH2 knock-down, a component of all Set1-like complexes in mammals that is required for H3K4me3 [14], and no splicing defect is observed in *cfp1*^{-/-} cells, specifically lacking the majority of H3K4me3 [62, 68, 83]. This might indicate an additional non-H3K4me3-dependent role of CHD1 in splicing. Other studies claim that H3K4me3 can regulate the alternative splicing of the *FGFR2* pre-mRNA between different cell types, with a variety of histone post-translation modifications including H3K4me1 and H3K4me3 differentially enriched or depleted over alternatively spliced exons [106]. Splicing is extremely sensitive to transcription elongation speed [107, 108] and many effects ascribed to H3K4me3 might simply reflect alterations to transcription elongation. For example, the dynamic acetylation of histones H3 and H4 in nucleosomes also containing H3K4me3 is proposed to affect the splicing of the *MCL1* pre-mRNA [109]. Interestingly, splicing can also influence H3K4me3, perhaps indicating a feedback mechanism for this process [110]. Thus, direct links between H3K4me3 and splicing are currently tenuous.

H3K4me3 and transcription termination linked to transcript degradation

A role for H3K4me3 has been described in the promotion of efficient Nrd1-dependent termination of RNA polymerase II transcription in yeast [111]. Nrd1, in combination with its complex members Nab3 and Sen1, binds to specific sites in the nascent RNA and mediates the termination of non-coding RNAs and some short mRNAs that are then targeted to the exosome for degradation or trimming [112]. Whilst there is reduced Nrd1 recruitment in the absence of H3K4 methylation [111], the mechanism for H3K4me3-stimulated

termination may also involve the ability of H3K4me3 to recruit both lysine acetyltransferases and deacetylases [113–115]. These control the levels of histone acetylation, which influence the kinetics of early elongation by RNA polymerase II and, therefore, may affect the efficiency of Nrd1-mediated termination. Thus the putative role of H3K4me3 in Nrd1-dependent termination may be an indirect consequence of altered transcription speed.

H3K4me3 is linked to transcriptional responsiveness

Although H3K4me3 increases over genes during their transcription induction, the modification often persists for many hours after transcription is shut down. This was shown in yeast at *GAL10*, where H3K4me3 remained on the chromatin for 5 h after inactivation of transcription and dissociation of Set1 [116]. In larger-scale yeast studies, the slowness of H3K4me3 to change was also shown for genes that were downregulated during the induction of meiosis [38] or upon diamide or osmotic stress [79, 117], especially compared to the rapid dynamic changes in levels of histone acetylation. Although this could just reflect the higher stability of H3K4me3 compared to histone acetylation, perhaps due to less efficient active removal by demethylation and/or reliance on a mechanism involving dilution of histones during replication [118], it was speculated that the maintenance of H3K4me3 might be advantageous in a reversal of conditions, allowing re-adjustment of the transcriptional response more rapidly [38]. In fact, high residual H3K4me3 has also been observed following transcription at trainable drought-resistance genes in *A. thaliana* and has been linked to faster transcriptional responses upon re-exposure to stress, a phenomenon known as transcriptional memory [119].

H3K4me3 promotes transcriptional consistency

Many genes can be expressed at very different levels in individual cells in a large population. This variation generates phenotypic heterogeneity and can be advantageous. For example, if not all cells in a population respond to a particular signal, the non-responders are primed for expression if conditions change again or return to pre-signalling, a phenomenon seen clearly in cell heterogeneity in cancer [120]. This variability is often referred to as gene expression noise. Methylation states at H3K4 are known to influence noise. In *D. discoideus*, H3K4 methylation can stabilise transcriptional frequency between generations, perhaps transmitting a memory of the previous state to the progeny [49]. In agreement, relative to acetylation of H3K9, H3K4me3 levels are higher at essential, periodically expressed yeast genes than non-essential, responsive genes, leading to the hypothesis that H3K4me3 is associated with expression stability in yeast [121]. Further support to the idea that H3K4me3 reduces variability in gene expression between cells comes from correlations between the presence of H3K4 methylation and gene expression noise for over 2,000 GFP-tagged yeast proteins [122]. H3K4me3 in the region of the +1 nucleosome

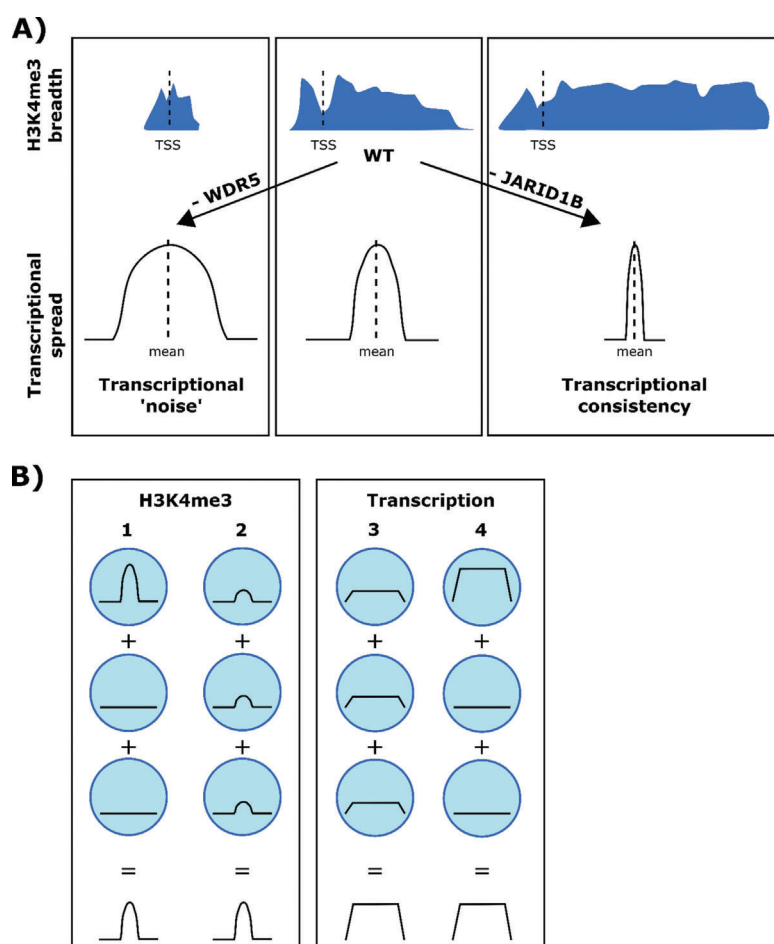


Figure 4. H3K4me3 influences events in single cells. **A:** Schematic showing the relationship between the breadth of the H3K4me3 peak in a population of cells and the variation in transcript levels in single cells (transcriptional spread) [4]. At some genes, knocking down WDR5 reduces H3K4me3 peak breadth and increases transcriptional noise. At other genes, knocking down JARID1B increases peak breadth and increases transcriptional consistency/reduces noise. Not all genes show changes in peak breadth upon knock-down of WDR5 or JARID1B. **B:** Schematic showing the theoretical relationship between peaks of H3K4me3 and levels of transcription in three individual cells (blue circles) and, below, the sum of these representing a population average. For example, if H3K4me3 is instructive for transcription then the cells in column 1 should show the transcription profile in column 4. If there is no relationship at the single-cell level, then the profile in column 1 might lead to the transcription profile in column 3. Note at the population level these different states cannot be distinguished and H3K4me3 appears linked to transcription.

(0–300 bp relative to the TSS) anti-correlates with noise ($r_s = -0.410$) while H3K4me1/2 correlate positively with noise ($r_s = +0.347$; $+0.296$, respectively). This reinforces the need to study each separate modification state in isolation, as loss of all methylation states in a *set1Δ* strain reduces noise at selected genes, probably reflecting the loss of H3K4me1/2, while loss of Spp1, and thus the majority of H3K4me3, increased noise (unpublished). These associations between H3K4me3 and transcriptional consistency are supported by a study of the *breadth* of the H3K4me3 peak in a range of cell

types and organisms (Fig. 4A) [4, 45]. The broadest peaks are found on the most essential genes for a particular cell type or key genes for basic cell functions across all cell types and correlate with transcriptional consistency. Variations in peak breadth are observed from yeast, flies, worms, plants to mammals but it remains to be seen whether breadth is an important determinant of consistency in many of these organisms. In mouse neural progenitor cells, however, knock-down of WDR5, a component of all the Set1/Set1-like complexes in mammals, decreases both H3K4me3 peak breadth and transcriptional consistency whereas depletion of one of the H3K4 demethylases JARID1B [123] increases both H3K4me3 peak breadth and transcriptional consistency at selected genes [4]. It should be noted that WDR5 is also present in complexes that contain the lysine acetyltransferases Gcn5 or MOF [124, 125] and so the effects of depletion on consistency may be indirect. Further experiments more specifically affecting H3K4me3 are required, but the opposing effects of WDR5 and JARID1B knock-downs on consistency are promising. On a population level, H3K4me3 may, therefore, not affect the mean level of transcripts or transcription, but could be influencing the spread of transcriptional responses between individual cells. This highlights the importance of considering single cell experiments to address questions such as whether H3K4me3 and transcription co-occur in the same cell (Fig. 4B). Most data on H3K4me3 *in vivo* come from ChIP experiments, where levels of the modification correlate with average transcription at the population level. However, ChIP levels say nothing about how many cells contain the H3K4me3 signal or whether those cells are undergoing transcription. For example, *PHO84* is a highly transcribed yeast gene with a broad peak of H3K4me3 [29], but RNA fluorescence *in situ* hybridisation reveals that *PHO84* transcripts are present only in 20–25% of the population [126]. Understanding how H3K4me3 is distributed in this population of cells would certainly clarify the relationship between H3K4me3 and transcripts.

Concluding remarks

Reassessment of the evidence that H3K4me3 influences transcription has yielded examples of both activation and repression of small subsets of genes. However, the overwhelming finding from most genome-wide transcript and transcription studies upon removal of the majority of H3K4me3 is that actually very few transcription events/mRNAs change at the population level, which is remarkable considering the strong correlations with transcripts and transcription. Rather than a major role in instructing transcription, H3K4me3 may instead be deposited as a result of transcription, influencing processes such as splicing, transcription termination, memory of previous states and transcriptional consistency. These events can alter the nature

and/or levels of protein produced in the population, without requiring a change to transcripts or transcription. One consideration is that because all studies are population-based, the consequences of the action of H3K4me3 may only be understood at the level of an individual cell. Levels of H3K4me3 are usually assessed using ChIP and reported as peaks at various sites in the genome. As yet we do not know how this reflects events, particularly transcription, in individual cells in the population and whether transcription and H3K4me3 will ever occur contemporaneously at an individual gene. A relatively new concept in our understanding of functions associated with H3K4me3 is peak breadth, whose size distribution, determined by the combined actions of the methyltransferases and demethylases, is remarkably conserved across eukaryotes despite large differences in gene size. Peak breadth is associated with both transcript levels and transcriptional consistency. Key questions include how breadth is determined, how H3K4me3 functions to suppress noise, whether H3K4me3 is antagonised by another modification, limiting its spread, and what is happening at the individual cell level, particularly during dynamically changing conditions. So H3K4me3, the universal mark associated with 'active' genes that every undergraduate with an interest in genes knows about, remains enigmatic.

Acknowledgments

The work was funded by The Wellcome Trust (WT089156MA and 097302/Z/11/Z), the BBSRC (BB/J001694/1), the EPA Cephalosporin Trust Fund (H.F) and the Medical Research Fund, University of Oxford (S.M). J.M holds stock in Oxford Biodynamics Ltd., Chronos Therapeutics Ltd., and Sibelius Ltd. but these holdings present no conflict of interest with work in this article.

The authors have declared no conflicts of interest.

References

- Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, et al. 2002. Active genes are tri-methylated at K4 of histone H3. *Nature* **419**: 407–11.
- Luger K, Mader AW, Richmond RK, Sargent DF, et al. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251–60.
- Howe FS, Boubriak I, Sale MJ, Nair A, et al. 2014. Lysine acetylation controls local protein conformation by influencing proline isomerization. *Mol Cell* **55**: 733–44.
- Benayoun BA, Pollina EA, Ucar D, Mahmoudi S, et al. 2014. H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell* **158**: 673–88.
- Ramakrishnan S, Pokhrel S, Palani S, Pflueger C, et al. 2016. Counteracting H3K4 methylation modulators Set1 and Jhd2 co-regulate chromatin dynamics and gene transcription. *Nat Commun* **7**: 11949.
- Shilatifard A. 2012. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem* **81**: 65–95.
- Yu BD, Hess JL, Horning SE, Brown GA, et al. 1995. Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378**: 505–8.
- Glaser S, Schaft J, Lubitz S, Vintersten K, et al. 2006. Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* **133**: 1423–32.
- Lee J, Saha PK, Yang Q-H, Lee S, et al. 2008. Targeted inactivation of MLL3 histone H3-Lys-4 methyltransferase activity in the mouse reveals vital roles for MLL3 in adipogenesis. *Proc Natl Acad Sci USA* **105**: 19229–34.
- Wu M, Wang PF, Lee J-S, Martin-Brown S, et al. 2008. Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Mol Cell Biol* **28**: 7337–44.
- Ardehali MB, Mei A, Zobeck KL, Caron M, et al. 2011. Drosophila Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription. *EMBO J* **30**: 2817–28.
- Mohan M, Herz H-M, Smith ER, Zhang Y, et al. 2011. The COMPASS family of H3K4 methylases in drosophila. *Mol Cell Biol* **31**: 4310–8.
- Dou Y, Milne TA, Ruthenburg AJ, Lee S, et al. 2006. Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol* **13**: 713–9.
- Steward MM, Lee J-S, O'Donovan A, Wyatt M, et al. 2006. Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nat Struct Mol Biol* **13**: 852–4.
- Wang P, Lin C, Smith E, Guo H, et al. 2009. Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. *Mol Cell Biol* **29**: 6074–85.
- Hu D, Gao X, Morgan MA, Herz H-M, et al. 2013. The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. *Mol Cell Biol* **33**: 4745–54.
- Rao RC, Dou Y. 2015. Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat Rev Cancer* **15**: 334–46.
- Greer EL, Maures TJ, Hauswirth AG, Green EM, et al. 2010. Members of the H3K4 trimethylation complex regulate lifespan in a germline-dependent manner in *C. elegans*. *Nature* **466**: 383–7.
- Ingham PW. 1998. Trithorax and the regulation of homeotic gene expression in Drosophila: a historical perspective. *Int J Dev Biol* **42**: 423–9.
- Wysocka J, Swigut T, Milne TA, Dou Y, et al. 2005. WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121**: 859–72.
- Ang YS, Tsai SY, Lee DF, Monk J, et al. 2011. Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* **145**: 183–7.
- Jiang H, Shukla A, Wang X, Chen WY, et al. 2011. Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains. *Cell* **144**: 513–25.
- Zhang H, Gayen S, Xiong J, Zhou B, et al. 2016. MLL1 inhibition reprograms epiblast stem cells to naive pluripotency. *Cell Stem Cell* **18**: 481–94.
- Bernstein BE, Humphrey EL, Erlich RL, Schneider R, et al. 2002. Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc Natl Acad Sci USA* **99**: 8695–700.
- Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, et al. 2005. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**: 169–81.
- Liu CL, Kaplan T, Kim M, Buratowski S, et al. 2005. Single-nucleosome mapping of histone modifications in *S. cerevisiae*. *PLoS Biol* **3**: e328.
- Barski A, Cuddapah S, Cui K, Roh TY, et al. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* **129**: 823–37.
- Guenther MG, Levine SS, Boyer LA, Jaenisch R, et al. 2007. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**: 77–88.
- Kirmizis A, Santos-Rosa H, Penkett CJ, Singer MA, et al. 2007. Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. *Nature* **449**: 928–32.
- Li X, Wang X, He K, Ma Y, et al. 2008. High-resolution mapping of epigenetic modifications of the rice genome uncovers interplay between DNA methylation, histone methylation, and gene expression. *Plant Cell* **20**: 259–76.
- Zhang X, Bernatavichute YV, Cokus S, Pellegrini M, et al. 2009. Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biol* **10**: R62.
- van Dijk K, Ding Y, Malkaram S, Riethoven J-JM, et al. 2010. Dynamic changes in genome-wide histone H3 lysine 4 methylation patterns in response to dehydration stress in *Arabidopsis thaliana*. *BMC Plant Biol* **10**: 238.
- Guillemette B, Drogaris P, Lin H-HS, Armstrong H, et al. 2011. H3 lysine 4 is acetylated at active gene promoters and is regulated by H3 lysine 4 methylation. *PLoS Genet* **7**: e1001354.
- Ram O, Goren A, Amit I, Shores N, et al. 2011. Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. *Cell* **147**: 1628–39.

35. **Chen K, Chen Z, Wu D, Zhang L**, et al. 2015. Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. *Nat Genet* **47**: 1149–57.
36. **Peña PV, Hom RA, Hung T, Lin H**, et al. 2008. Histone H3K4me3 binding is required for the DNA repair and apoptotic activities of ING1 tumor suppressor. *J Mol Biol* **380**: 303–12.
37. **Faucher D, Wellinger RJ**. 2010. Methylated H3K4, a transcription-associated histone modification, is involved in the DNA damage response pathway. *PLoS Genet* **6**: e1001082.
38. **Borde V, Robine N, Lin W, Bonfils S**, et al. 2009. Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites. *EMBO J* **28**: 99–111.
39. **Acquaviva L, Székelyölyi L, Dichtl B, Dichtl BS**, et al. 2013. The COMPASS subunit Spp1 links histone methylation to initiation of meiotic recombination. *Science* **339**: 215–8.
40. **Sommermeier V, Béneut C, Chaplais E, Serrentino ME**, et al. 2013. Spp1, a member of the Set1 Complex, promotes meiotic DSB formation in promoters by tethering histone H3K4 methylation sites to chromosome axes. *Mol Cell* **49**: 43–54.
41. **Quinodoz M, Gobet C, Naef F, Gustafson KB**. 2014. Characteristic bimodal profiles of RNA polymerase II at thousands of active mammalian promoters. *Genome Biol* **15**: R85.
42. **Bornelöv S, Komorowski J, Wadelius C**. 2015. Different distribution of histone modifications in genes with unidirectional and bidirectional transcription and a role of CTCF and cohesin in directing transcription. *BMC Genomics* **16**: 300.
43. **Margaritis T, Oreal V, Brabers NACH, Maestroni L**, et al. 2012. Two distinct repressive mechanisms for histone 3 lysine 4 methylation through promoting 3'-end antisense transcription. *PLoS Genet* **8**: e1002952.
44. **Lavender CA, Cannady KR, Hoffman JA, Trotter KW**, et al. 2016. Downstream antisense transcription predicts genomic features that define the specific chromatin environment at mammalian promoters. *PLoS Genet* **12**: e1006224.
45. **Benayoun BA, Pollina EA, Ucar D, Mahmoudi S**, et al. 2014. H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell* **163**: 1281–6.
46. **Zhang K, Lin W, Latham JA, Riefler GM**, et al. 2005. The Set1 methyltransferase opposes Ipl1 aurora kinase functions in chromosome segregation. *Cell* **122**: 723–34.
47. **Ding Y, Ndamukong I, Xu Z, Lapko H**, et al. 2012. ATX1-generated H3K4me3 is required for efficient elongation of transcription, not initiation, at ATX1-regulated genes. *PLoS Genet* **8**: e1003111.
48. **Mann RK, Grunstein M**. 1992. Histone H3 N-terminal mutations allow hyperactivation of the yeast GAL1 gene in vivo. *EMBO J* **11**: 3297–306.
49. **Muramoto T, Müller I, Thomas G, Melvin A**, et al. 2010. Methylation of H3K4 is required for inheritance of active transcriptional states. *Curr Biol* **20**: 397–406.
50. **Morillon A, Karabetsou N, Nair A, Mellor J**. 2005. Dynamic lysine methylation on histone H3 defines the regulatory phase of gene transcription. *Mol Cell* **18**: 723–34.
51. **Schneider J, Wood A, Lee J-S, Schuster R**, et al. 2005. Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. *Mol Cell* **19**: 849–56.
52. **Thomson JP, Skene PJ, Selfridge J, Clouaire T**, et al. 2010. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* **464**: 1082–6.
53. **Walter D, Matter A, Fahrenkrog B**. 2014. Loss of histone H3 methylation at lysine 4 triggers apoptosis in *Saccharomyces cerevisiae*. *PLoS Genet* **10**: e1004095.
54. **Carlone DL, Skalik DG**. 2001. CpG binding protein is crucial for early embryonic development. *Mol Cell Biol* **21**: 7601–6.
55. **Carlone DL, Lee J-H, Young SRL, Dobrota E**, et al. 2005. Reduced genomic cytosine methylation and defective cellular differentiation in embryonic stem cells lacking CpG binding protein. *Mol Cell Biol* **25**: 4881–91.
56. **Deaton A, Bird A**. 2011. CpG islands and the regulation of transcription. *Genes Dev* **25**: 1010–22.
57. **Voo KS, Carlone DL, Jacobsen BM, Flodin A**, et al. 2000. 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**: 2108–21.
58. **Eberl HC, Spruijt CG, Kelstrup CD, Vermeulen M**, et al. 2013. A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics. *Mol Cell* **49**: 368–78.
59. **Birke M, Schreiner S, García-Cuellar M-P, Mahr K**, et al. 2002. The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic Acids Res* **30**: 958–65.
60. **Bach C, Mueller D, Buhl S, Garcia-Cuellar M-P**, et al. 2009. Alterations of the CxxC domain preclude oncogenic activation of mixed-lineage leukemia 2. *Oncogene* **28**: 815–23.
61. **Wang Z, Song J, Milne TA, Wang GG**, et al. 2010. Pro isomerization in MLL1 PHD3-Bromo cassette connects H3K4me3 readout to Cyp33 and HDAC-mediated repression. *Cell* **141**: 1183–94.
62. **Clouaire T, Webb S, Skene P, Illingworth R**, et al. 2012. Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. *Genes Dev* **26**: 1714–28.
63. **Hu D, Garruss AS, Gao X, Morgan MA**, et al. 2013. The Mll2 branch of the COMPASS family regulates bivalent promoters in mouse embryonic stem cells. *Nat Struct Mol Biol* **29**: 1093–7.
64. **Denissov S, Hofmeister H, Marks H, Kranz A**, et al. 2014. Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. *Development* **141**: 526–37.
65. **Shi X, Kachirskaja I, Walter KL, Kuo J-HA**, et al. 2007. Proteome-wide analysis in *Saccharomyces cerevisiae* identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. *J Biol Chem* **282**: 2450–5.
66. **Takahashi Y-H, Westfield GH, Oleskie AN, Trievel RC**, et al. 2011. Structural analysis of the core COMPASS family of histone H3K4 methylases from yeast to human. *Proc Natl Acad Sci USA* **108**: 20526–31.
67. **Kim J, Kim J-A, McGinty RK, Nguyen UTT**, et al. 2013. The n-SET domain of Set1 regulates H2B ubiquitylation-dependent H3K4 methylation. *Mol Cell* **49**: 1121–33.
68. **Clouaire T, Webb S, Bird A**. 2014. Cfp1 is required for gene expression-dependent H3K4 trimethylation and H3K9 acetylation in embryonic stem cells. *Genome Biol* **15**: 451.
69. **Lenstra TL, Benschop JJ, Kim T, Schulze JM**, et al. 2011. The specificity and topology of chromatin interaction pathways in yeast. *Mol Cell* **42**: 536–49.
70. **Churchman LS, Weissman JS**. 2011. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **469**: 368–73.
71. **Nguyen T, Fischl H, Howe FS, Woloszczuk R**, et al. 2014. Transcription mediated insulation and interference direct gene cluster expression switches. *Elife* **3**: 1–21.
72. **Ingvarsdottir K, Edwards C, Lee MG, Lee J-S**, et al. 2007. Histone H3 K4 demethylation during activation and attenuation of GAL1 transcription in *Saccharomyces cerevisiae*. *Mol Cell Biol* **27**: 7856–64.
73. **Liang G, Klose RJ, Gardner KE, Zhang Y**. 2007. Yeast Jhd2p is a histone H3 Lys4 trimethyl demethylase. *Nat Struct Mol Biol* **14**: 243–5.
74. **Seward DJ, Cubberley G, Kim S, Schonewald M**, et al. 2007. Demethylation of trimethylated histone H3 Lys4 in vivo by JARID1 JmjC proteins. *Nat Struct Mol Biol* **14**: 240–2.
75. **Tu S, Bulloch EMM, Yang L, Ren C**, et al. 2007. Identification of histone demethylases in *Saccharomyces cerevisiae*. *J Biol Chem* **282**: 14262–71.
76. **Huang F, Chandrasekharan MB, Chen Y-C, Bhaskara S**, et al. 2010. The JmjN domain of Jhd2 is important for its protein stability, and the plant homeodomain (PHD) finger mediates its chromatin association independent of H3K4 methylation. *J Biol Chem* **285**: 24548–61.
77. **Cano-Rodriguez D, Gjaltema RAF, Jilderda LJ, Jellema P**, et al. 2016. Writing of H3K4Me3 overcomes epigenetic silencing in a sustained, but context-dependent manner. *Nat Commun* **7**: 1–11.
78. **Weiner A, Chen HV, Liu CL, Rahat A**, et al. 2012. Systematic dissection of roles for chromatin regulators in a yeast stress response. *PLoS Biol* **10**: e1001369.
79. **Weiner A, Hsieh T-HS, Appleboim A, Chen HV**, et al. 2015. High-resolution chromatin dynamics during a yeast stress response. *Mol Cell* **58**: 371–86.
80. **Ruthenburg AJ, Allis CD, Wysocka J**. 2007. Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell* **25**: 15–30.
81. **Ruthenburg AJ, Li H, Milne TA, Dewell S**, et al. 2011. Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell* **145**: 692–706.
82. **Santos-Rosa H, Schneider R, Bernstein BE, Karabetsou N**, et al. 2003. Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. *Mol Cell* **12**: 1325–32.
83. **Sims RJ, Millhouse S, Chen CF, Lewis BA**, et al. 2007. Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of

- transcription postinitiation factors and pre-mRNA splicing. *Mol Cell* **28**: 665–76.
84. Pinskaya M, Nair A, Clynes D, Morillon A, et al. 2009. Nucleosome remodeling and transcriptional repression are distinct functions of Isw1 in *Saccharomyces cerevisiae*. *Mol Cell Biol* **29**: 2419–30.
 85. Lauberth SM, Nakayama T, Wu X, Ferris AL, et al. 2013. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**: 1021–36.
 86. Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, et al. 2007. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**: 58–69.
 87. Gangloff Y-G, Pointud J-C, Thuault S, Carré L, et al. 2001. The TFIID components human TAFII 140 and Drosophila BIP2 (TAFII 155) are novel metazoan homologues of yeast TAFII 47 containing a histone fold and a PHD finger. *Mol Cell Biol* **21**: 5109–21.
 88. Lawit SJ, O'Grady K, Gurley WB, Czarnecka-Verner E. 2007. Yeast two-hybrid map of Arabidopsis TFIID. *Plant Mol Biol* **64**: 73–87.
 89. Pavri R, Zhu B, Li G, Trojer P, et al. 2006. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125**: 703–17.
 90. Krogan NJ, Dover J, Wood A, Schneider J, et al. 2003. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell* **11**: 721–9.
 91. Smith ST, Petruk S, Sedkov Y, Cho E, et al. 2004. Modulation of heat shock gene expression by the TAC1 chromatin-modifying complex. *Nat Cell Biol* **6**: 162–7.
 92. Milne TA, Dou Y, Martin ME, Brock HW, et al. 2005. MLL associates specifically with a subset of transcriptionally active target genes. *Proc Natl Acad Sci USA* **102**: 14765–70.
 93. Lee J-H, Skalik DG. 2008. Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. *Mol Cell Biol* **28**: 609–18.
 94. Ding Y, Avramova Z, Fromm M. 2011. Two distinct roles of ARABIDOPSIS HOMOLOG OF TRITHORAX1 (ATX1) at promoters and within transcribed regions of ATX1-regulated genes. *Plant Cell* **23**: 350–63.
 95. Yang Y, Li W, Hoque M, Hou L, et al. 2016. PAF complex plays novel subunit-specific roles in alternative cleavage and polyadenylation. *PLoS Genet* **12**: e1005794.
 96. Osley MA. 2006. Regulation of histone H2A and H2B ubiquitylation. *Briefings Funct Genomics Proteomics* **5**: 179–89.
 97. Dover J, Schneider J, Tawiah-Boateng MA, Wood A, et al. 2002. Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *J Biol Chem* **277**: 28368–71.
 98. Sun Z-W, Allis CD. 2002. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418**: 104–8.
 99. Huang F, Ramakrishnan S, Pokhrel S, Pflueger C, et al. 2015. Interaction of the Jhd2 histone H3 Lys-4 demethylase with chromatin is controlled by histone H2A surfaces and restricted by H2B ubiquitination. *J Biol Chem* **290**: 28760–77.
 100. Kuang Z, Cai L, Zhang X, Ji H, et al. 2014. High-temporal-resolution view of transcription and chromatin states across distinct metabolic states in budding yeast. *Nat Struct Mol Biol* **21**: 854–63.
 101. Sollier J, Lin W, Soustelle C, Suhre K, et al. 2004. Set1 is required for meiotic S-phase onset, double-strand break formation and middle gene expression. *EMBO J* **23**: 1957–67.
 102. Le Martelot G, Canella D, Symul L, Migliavacca E, et al. 2012. Genome-wide RNA polymerase II profiles and RNA accumulation reveal kinetics of transcription and associated epigenetic changes during diurnal cycles. *PLoS Biol* **10**: e1001442.
 103. Vastenhouw NL, Zhang Y, Woods IG, Imam F, et al. 2010. Chromatin signature of embryonic pluripotency is established during genome activation. *Nature* **464**: 922–6.
 104. Hontelez S, van Kruisbergen I, Georgiou G, van Heeringen SJ, et al. 2015. Embryonic transcription is controlled by maternally defined chromatin state. *Nat Commun* **6**: 10148.
 105. Sims RJ, Chen C-F, Santos-Rosa H, Kouzarides T, et al. 2005. Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *J Biol Chem* **280**: 41789–92.
 106. Luco RF, Pan Q, Tominaga K, Blencowe BJ, et al. 2010. Regulation of alternative splicing by histone modifications. *Science* **327**: 996–1000.
 107. de la Mata M, Alonso CR, Kadener S, Fededa JP, et al. 2003. A slow RNA polymerase II affects alternative splicing in vivo. *Mol Cell* **12**: 525–32.
 108. Naftelberg S, Schor IE, Ast G, Kornblihtt AR. 2015. Regulation of alternative splicing through coupling with transcription and chromatin structure. *Annu Rev Biochem* **84**: 165–98.
 109. Khan DH, Gonzalez C, Tailor N, Hamedani MK, et al. 2016. Dynamic histone acetylation of H3K4me3 nucleosome regulates MCL1 pre-mRNA splicing. *J Cell Physiol* **9999**: 1–9.
 110. Bieberstein NI, Oesterreich FC, Straube K, Neugebauer KM. 2012. First exon length controls active chromatin signatures and transcription. *Cell Rep* **2**: 62–8.
 111. Terzi N, Churchman LS, Vasiljeva L, Weissman JS, et al. 2011. H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway. *Mol Cell Biol* **31**: 3569–83.
 112. Porrua O, Libri D. 2015. Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat Rev Mol Cell Biol* **16**: 190–202.
 113. Wang S-S, Zhou BO, Zhou J-Q. 2011. Histone H3 lysine 4 hypermethylation prevents aberrant nucleosome remodeling at the PHO5 promoter. *Mol Cell Biol* **31**: 3171–81.
 114. Martin DGE, Baetz K, Shi X, Walter KL, et al. 2006. The Yng1p plant homeodomain finger is a methyl-histone binding module that recognizes lysine 4-methylated histone H3. *Mol Cell Biol* **26**: 7871–9.
 115. Taverna SD, Ilin S, Rogers RS, Tanny JC, et al. 2006. Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Mol Cell* **24**: 785–96.
 116. Ng HH, Robert F, Young RA, Struhl K. 2003. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* **11**: 709–19.
 117. Magraner-Pardo L, Pelechano V, Coloma MD, Tordera V. 2014. Dynamic remodeling of histone modifications in response to osmotic stress in *Saccharomyces cerevisiae*. *BMC Genomics* **15**: 247.
 118. Radman-Livaja M, Liu CL, Friedman N, Schreiber SL, et al. 2010. Replication and active demethylation represent partially overlapping mechanisms for erasure of H3K4me3 in budding yeast. *PLoS Genet* **6**: e1000837.
 119. Ding Y, Fromm M, Avramova Z. 2012. Multiple exposures to drought “train” transcriptional responses in Arabidopsis. *Nat Commun* **3**: 740.
 120. Brock A, Krause S, Ingber DE. 2015. Control of cancer formation by intrinsic genetic noise and microenvironmental cues. *Nat Rev Cancer* **15**: 499–509.
 121. Mosesson Y, Voichek Y, Barkai N. 2014. Divergence and selectivity of expression-coupled histone modifications in budding yeasts. *PLoS ONE* **9**: e01538.
 122. Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, et al. 2006. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441**: 840–6.
 123. Kooistra SM, Helin K. 2012. Molecular mechanisms and potential functions of histone demethylases. *Nat Publ Gr* **13**: 297–311.
 124. Dou Y, Milne TA, Tackett AJ, Smith ER, et al. 2005. Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* **121**: 873–85.
 125. Wang YL, Faiola F, Xu M, Pan S, et al. 2008. Human ATAC is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold module that interacts with the TATA-binding protein. *J Biol Chem* **283**: 33808–15.
 126. Castelnovo M, Rahman S, Guffanti E, Infantino V, et al. 2013. Bimodal expression of PHO84 is modulated by early termination of antisense transcription. *Nat Struct Mol Biol* **20**: 851–8.
 127. Booth GT, Wang IX, Cheung VG, Lis JT. 2016. Divergence of a conserved elongation factor and transcription regulation in budding and fission yeast. *Genome Res* **26**: 799–811.
 128. Vermeulen M, Eberl HC, Matarese F, Marks H, et al. 2010. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* **142**: 967–80.
 129. Bian C, Xu C, Ruan J, Lee KK, et al. 2011. Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. *EMBO J* **30**: 2829–42.
 130. Wen H, Li J, Song T, Lu M, et al. 2010. Recognition of histone H3K4 trimethylation by the plant homeodomain of PHF2 modulates histone demethylation. *J Biol Chem* **285**: 9322–6.
 131. Feng W, Yonezawa M, Ye J, Jenuwein T, et al. 2010. PHF8 activates transcription of rRNA genes through H3K4me3 binding and H3K9me1/2 demethylation. *Nat Struct Mol Biol* **17**: 445–50.
 132. Fortschegger K, de Graaf P, Outchkourov NS, van Schaik FM, et al. 2010. PHF8 targets histone methylation and RNA polymerase II to activate transcription. *Mol Cell Biol* **30**: 3286–98.
 133. Chung H-R, Xu C, Fuchs A, Mund A, et al. 2016. PHF13 is a molecular reader and transcriptional co-regulator of H3K4me2/3. *Elife* **5**: e10607.

134. **Shi X, Hong T, Walter KL, Ewalt M**, et al. 2006. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442**: 96–9.
135. **Flanagan JF, Mi L-Z, Chruszcz M, Cymborowski M**, et al. 2005. Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* **438**: 1181–5.
136. **Li H, Ilin S, Wang W, Duncan EM**, et al. 2006. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442**: 91–5.
137. **Wysocka J, Swigut T, Xiao H, Milne TA**, et al. 2006. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**: 86–90.
138. **Kusch T, Mei A, Nguyen C**. 2014. Histone H3 lysine 4 trimethylation regulates cotranscriptional H2A variant exchange by Tip60 complexes to maximize gene expression. *Proc Natl Acad Sci USA* **111**: 4850–5.
139. **Nojima T, Gomes T, Grosso ARF, Kimura H**, et al. 2015. Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell* **161**: 526–40.