



# Shaping the cellular landscape with Set2/SETD2 methylation

Stephen L. McDaniel<sup>1</sup> · Brian D. Strahl<sup>2,3,4</sup>

Received: 1 February 2017 / Revised: 24 March 2017 / Accepted: 28 March 2017 / Published online: 6 April 2017  
© Springer International Publishing 2017

**Abstract** Chromatin structure is a major barrier to gene transcription that must be disrupted and re-set during each round of transcription. Central to this process is the Set2/SETD2 methyltransferase that mediates co-transcriptional methylation to histone H3 at lysine 36 (H3K36me). Studies reveal that H3K36me not only prevents inappropriate transcriptional initiation from arising within gene bodies, but that it has other conserved functions that include the repair of damaged DNA and regulation of pre-mRNA splicing. Consistent with the importance of Set2/SETD2 in chromatin biology, mutations of SETD2, or mutations at or near H3K36 in H3.3, have recently been found to underlie cancer development. This review will summarize the latest insights into the functions of Set2/SETD2 in genome regulation and cancer development.

**Keywords** Histone · H3K36 methylation · Set2 · SETD2 · Transcriptional regulation · Cryptic transcription

## Introduction

A major mechanism by which eukaryotic cells regulate gene transcription is through the packaging of their genomes into chromatin—a complex of DNA and histone/non-histone proteins that serves as the “gate keeper” of genetic information. The fundamental unit of chromatin is the nucleosome, which contains ~147 base pairs of DNA wrapped around eight histone proteins: two each of H2A, H2B, H3, and H4 [1–4]. In addition to the actions of histone variants and specialized protein machines that assemble, disassemble, and move nucleosomes along DNA, histone post-translational modifications (PTMs) play critical roles in the accessing and regulation of our genome [5–7]. In particular, histone lysine methylation is known to direct the organization of basic chromatin states such as euchromatin and heterochromatin, as well as the precise organization of chromatin across genes before, during, and after the process of transcription [8]. In this review, we highlight recent findings pertaining to the function of the histone H3 at lysine 36 (H3K36) methyltransferase Set2/SETD2, and outline future questions that need to be addressed to fully understand the multi-faceted roles of this enzyme in cell biology.

## Histone lysine methylation in transcription

Lysine residues can be mono-, di-, or tri-methylated retaining the positive charge of the residue regardless of the number of methyl groups added. Lysine methylation is associated with both active and repressed states of transcription, primarily depending on which lysine residue is methylated and in which methylation state (mono-, di-, or tri-) that residue exists [8]. In general, lysine methylation functions

✉ Brian D. Strahl  
brian\_strahl@med.unc.edu

<sup>1</sup> Department of Biomolecular Chemistry, University of Wisconsin, Madison, WI 53706, USA

<sup>2</sup> Curriculum in Genetics and Molecular Biology, University of North Carolina, Chapel Hill, NC 27599, USA

<sup>3</sup> Department of Biochemistry, University of North Carolina, Chapel Hill, NC 27599, USA

<sup>4</sup> Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599, USA

to recruit effector proteins bearing specialized domains that uniquely recognize these modified sites. Nevertheless, the presence of some methylated lysine residues in the nucleosome core close to where DNA wraps or near the dyad axis close to where DNA enters and exists the nucleosome (e.g., H3K56 methylation), suggests that some methylation events function, in part, to fine-tune or tweak nucleosome-DNA interactions [9].

In the budding yeast *Saccharomyces cerevisiae*, there are three well studied methylated lysine residues: histone H3 lysine 4 methylation (H3K4me), H3K36me, and histone H3 lysine 79 methylation (H3K79me). All three residues and their methylation events are conserved from yeast to humans, making budding yeast an attractive model system to study the fundamental biology of these marks. Although all three marks function in transcriptional regulation in yeast and mammalian cells [10], their precise molecular functions in transcription and other biological processes remain poorly understood. Studies have generally found that H3K4me<sub>2/3</sub> is located primarily at promoter regions (note that enhancers in metazoans are marked with H3K4me<sub>1</sub>) whereas H3K36me<sub>2/3</sub> is localized towards the 3' regions of gene bodies [11, 12]. Interestingly, H3K79me<sub>2/3</sub> is generally uniform across actively transcribed regions of the genome in yeast but is largely found in promoter regions in metazoans.

## Set2 and H3K36 methylation

In *S. cerevisiae*, Set2 is the sole histone methyltransferase that acts on H3K36 [13]. Conversely, in higher eukaryotes, there are multiple H3K36 methyltransferases that are responsible for this mark. For example, in *Drosophila melanogaster*, dMes-4 is responsible for H3K36me<sub>1</sub> and H3K36me<sub>2</sub>, whereas dSet2 is primarily responsible for H3K36me<sub>3</sub> [14]. In humans, NSD1, NSD2, NSD3, SETMAR, SMYD2 and ASH1L are responsible for mediating H3K36me<sub>1</sub> and H3K36me<sub>2</sub> [15, 16], and SETD2, the human Set2 homolog, mediates H3K36me<sub>3</sub> [17]. Because Set2 is responsible for all H3K36me in yeast, and strains in which *SET2* has been deleted are viable, yeast has become a premiere model system for studying the function of H3K36me.

## The domain structure of Set2

### The SET domain

Set2 has several conserved domains (Fig. 1), the first being the SET domain itself, which consists of the AWS (associated with SET), SET, and PS (post-SET) motifs. The SET

domain is the catalytic domain of Set2 that performs the H3K36me transfer. Interestingly, although full-length Set2 prefers to methylate nucleosomal substrates, the isolated SET domain can also methylate free histones [18]. Thus, the C-terminus of Set2 is critical for directing substrate specificity (discussed below). In budding yeast, the SET domain is located at the extreme N-terminus of the protein, which differs from its location in higher eukaryotic Set2 homologs. SETD2, the human homolog of Set2, has an extended and low complexity N-terminus with no conserved domains and no known function [19]. This extended N-terminus is common in non-yeast eukaryotes and likely plays an important, albeit poorly understood regulatory role in methylation of histone and non-histone targets.

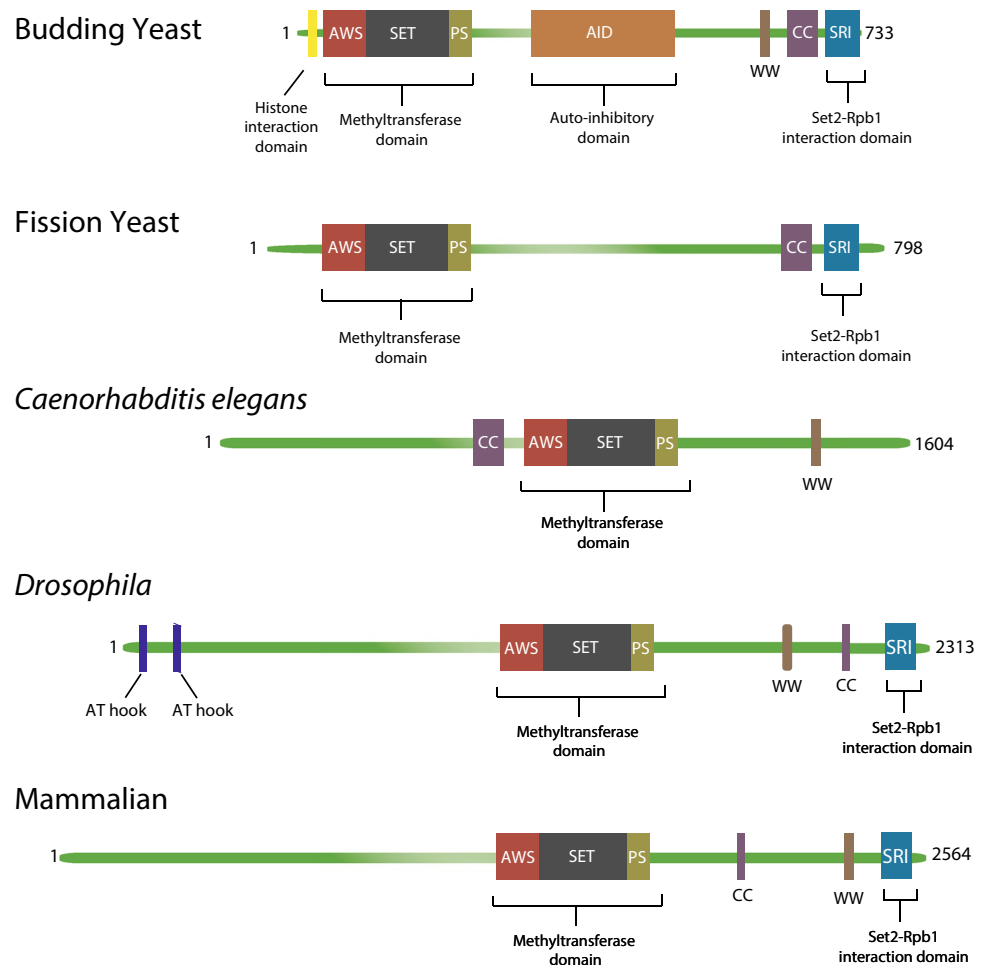
### The H4/H2A interaction domain

A histone H4/H2A interaction domain precedes the AWS-SET-PS domain in yeast Set2 [20, 21]. This domain, consisting of a short stretch of acidic residues, binds a basic region or patch located on the nucleosomal surface that encompasses residues located in H4 and H2A. Notably, disrupting the acidic residues in Set2, or eliminating the basic residues in H2A or H4 that create the Set2 binding surface, significantly reduces H3K36 methylation in vivo, thus defining the importance of this interaction for Set2 methylation in chromatin. Although further work is needed to clarify precisely how this nucleosomal surface interaction regulates H3K36 methylation, it is likely that this basic region in H4/H2A stabilizes Set2 on the nucleosome or positions the enzymatic domain over the H3K36 residue. Given the high conservation of Set2, it is likely that Set2 homologs in more complex organisms also retain the H4/H2A-interaction domain, but this conjecture remains to be tested.

### The WW and coiled-coiled domains

The SET domain is followed by two protein-protein interaction domains, a WW and a coiled-coiled (CC) domain. The WW domain, named for two conserved tryptophan residues, binds to proline rich proteins [22]. In the nucleus, proteins with this domain regulate transcription by binding to the C-terminal domain (CTD) of RNA polymerase II (RNAPII) [23]. Although the WW domain is conserved in Set2 homologs, its function, if any, and its binding partners are currently unknown. Notably, a precise deletion of the WW domain in yeast does not affect the association of Set2 with RNAPII or alter H3K36 methylation [24], thus we speculate that the WW domain participates in the association of Set2 with proteins that contribute to methylation of non-histone substrates (see below). In humans, the WW domain of SETD2 mediates an interaction of SETD2 with

**Fig. 1** Domain architecture of Set2/SETD2 homologs. Set2/SETD2-related enzymes share a highly conserved domain organization that encompasses a well-conserved catalytic region containing an Associated with SET (AWS), SET and post-SET (PS) domain, a WW and coiled-coil (CC) domain, followed by the Set2–Rpb1 interaction (SRI) domain. In *S. cerevisiae*, studies into the function of Set2 have revealed additional domains including an auto-inhibitory (AID) domain that regulates the activity of the SET domain and a H4/H2A interaction domain. Whether these additional domains are conserved outside of budding yeast is not yet known



the polyglutamine expansion region of the huntingtin (Htt) protein [25, 26]; in the expanded state, Htt leads to Huntington's disease. Interestingly, SETD2 was first identified through yeast two-hybrid screens (first named two-hybrid protein B or HYPB) [26]. However, aside from Htt interaction, we do not know the function of the WW domain in SETD2.

The CC domain in Set2, which follows the WW domain, is also a conserved protein–protein interaction motif, that, in certain proteins, promotes homo-dimerization [27]. This domain is also highly conserved within the family of Set2 enzymes, thus future work is needed to understand how it precisely functions. We speculate, given its close proximity, that the CC domain contributes to a recently described auto-inhibitory domain [18] that controls Set2 methylation output (see below).

### The auto-inhibitory domain (AID)

As mentioned above, recent studies show that Set2 contains an auto-inhibitory domain (AID) in the region between the SET domain and the WW domain [18]. Deletion of the

AID results in a hyperactive Set2, increasing the amount of H3K36me on nucleosomes and at genes. Further, the AID influences Set2 interactions RNAPII, as AID mutant versions of Set2 are able to bind to the unphosphorylated CTD of RNAPII, whereas wild-type Set2 can bind to only the hyperphosphorylated form of the CTD. Further work is needed to ascertain whether the AID is conserved in higher eukaryotes and how it regulates Set2 methylation and CTD interaction.

### The Set2–Rpb1 interacting (SRI) domain

Finally, Set2 contains a Set2–Rpb1 interacting (SRI) domain. This domain, which is highly conserved from yeast to humans (Fig. 1) [28], binds to serine 2 (Ser2) and serine 5 (Ser5) doubly phosphorylated CTD repeats of RNAPII [29–33]. Specifically, the SRI domain recognizes these phosphorylated residues across two heptapeptide repeats of the CTD [31, 32]. Intriguingly, the *Caenorhabditis elegans* Set2 homolog, MET-1, does not appear to contain a SRI domain based on sequence analysis in Pfam (<http://pfam.xfam.org>). Whether a non-canonical SRI domain exists in

this protein, and if it functions in conjunction with elongating RNAPII, awaits future testing.

Importantly, deletion of the SRI domain does not abolish chromatin localization of Set2, suggesting that additional mechanisms are primarily responsible for recruiting Set2 to chromatin [34]. Furthermore, ectopically expressed Set2 truncations lacking the SRI domain still catalyze H3K36me1 and limited amounts of H3K36me2, but they are unable to catalyze H3K36me3 [29]. One explanation for this result may be the ability of the SRI domain to regulate the AID of Set2 that itself antagonizes the SET domain [18]. In *Schizosaccharomyces pombe*, however, the SRI domain is completely dispensable for H3K36me2, but is required for H3K36me3 and plays a role in post-transcriptional gene silencing at subtelomeric regions of the genome [35]. However, mutants of Spt6 and Ctk1 that prevent CTD phosphorylation at Ser2 render chromatin incapable of being trimethylated by full-length recombinant Set2 [34]. Thus, it appears that Set2–CTD interaction is essential in some way for Set2 to catalyze nucleosomal H3K36me3, but, at the same time, the SRI domain also regulates the auto-inhibitory functions of Set2 [18]. Future experiments are needed to better define how the SRI–CTD interaction, and functions of the SRI domain that lie outside of CTD interaction, regulate the methylation states of Set2.

## Regulation of Set2 and H3K36 methylation

### Ser2 and Ser5 CTD phosphorylation of RNAPII

CTD phosphorylation plays a critical role in the regulation of H3K36 methylation. Early studies in yeast found that deletion of the major kinase complex that mediates Ser2 CTD phosphorylation, CTDK-1 (CDK12 in metazoans), results in an ablation of H3K36 methylation [24]. Consistent with this finding, deletions of the CTD or mutations of Ser2 residues (but not Ser5) in the CTD also caused a loss of H3K36 methylation [24, 33]. Intriguingly, examination of the interaction of the SRI domain with the CTD revealed that, although Ser2 phosphorylation is critical for Set2 binding, Set2 actually prefers to bind to doubly modified (Ser2 and Ser5 phosphorylated) CTD repeats—a phosphorylation pattern that is, in fact, mediated by the Ctk1 kinase of the CTDK-1 complex [29, 36]. However, in yeast, doubly modified Ser2 and Ser5 phosphorylated CTD repeat patterns appear to exist at very low levels [37], suggesting that Ser2 CTD phosphorylation may be the primary mark available for Set2 interaction.

In addition to Ctk1, the Bur1 kinase complex (Cdk9 of p-TEFb complex in metazoans) mediates Ser2 CTD phosphorylation in vitro and in vivo [38, 39]. Consistent with a role of Ser2 CTD phosphorylation in regulating

H3K36me, *bur1* temperature sensitive alleles, or a deletion of *BUR2* (the Bur1 cyclin), reduce H3K36me3 levels in yeast [40, 41]. Unlike the loss of Ctk1, loss of Bur1 activity does not significantly alter Ser2 and Ser5 phosphorylation levels in vivo [39, 42], suggesting that the impact of Bur1 on H3K36 methylation is largely independent of the CTD. The ability of Bur1 to affect H3K36 methylation is likely an indirect effect of Bur1's ability to regulate Spt4/5 (DSIF) and recruit the Paf1 transcription elongation complex (Paf1C) to genes, which then regulates Ser2 CTD phosphorylation [38, 39, 43–45].

An unexpected finding regarding the regulation of Set2 methylation by the CTD was the realization that Set2–RNAPII interaction is critical for Set2 stability. Loss of Ctk1, or mutation of elongation machinery that impacts Ser2 CTD phosphorylation (e.g., loss of PAF or Bur1 activity), results in a rapid turnover of Set2 [46]. These results indicate that a key function of CTD binding is to protect Set2 from degradation. Interestingly, removal of the SRI domain stabilizes Set2 in addition to uncoupling Set2 from the CTD. Thus, the SRI domain, which protects Set2 from degradation when bound to the CTD, must also contain a degradation signal that becomes exposed when Set2 is not CTD-bound [46]. Although the significance of Set2 turnover is not understood, this regulation is clearly important because a truncated form of Set2 is lethal when expressed in cells deficient in H2A.Z deposition [46]. The rapid turnover of Set2 appears to be conserved in humans, as a recent study identified SPOP, a key subunit of the CUL3 ligase complex, as a binding partner for SETD2 that mediates its turnover by the proteasome [47].

### The Spt6 histone chaperone

In addition to regulation by the CTD, a number of transcription elongation factors and histone chaperones also regulate Set2-mediated H3K36me. One such factor is Spt6, a highly conserved H3/H4 histone chaperone that also binds to the phosphorylated CTD of RNAPII to facilitate transcription elongation and nucleosome re-assembly in the wake of RNAPII elongation [48–53]. Mutations in multiple regions of Spt6 result in a reduction in H3K36me2 and/or H3K36me3 levels [34, 45, 54, 55]. Similar to Ctk1, mutations in Spt6 also reduce Set2 protein stability [34, 45]. Interestingly, in budding yeast, Ctk1 and Spt6 mutually regulate each other's protein stability [45], forming a feed-forward loop that maintains CTD Ser2 phosphorylation, and consequently, Set2 levels. The ability of Spt6 to maintain Ser2 CTD phosphorylation levels is likely behind the ability of Spt6 to control H3K36me. In human cells, however, the ability of Spt6 to regulate Set2 may be more

direct. Iws1 (Spn1 in yeast), which is a binding partner of Spt6 that recruits this chaperone to genes [53, 56], associates with human SEDT2 and recruits it to genes [53].

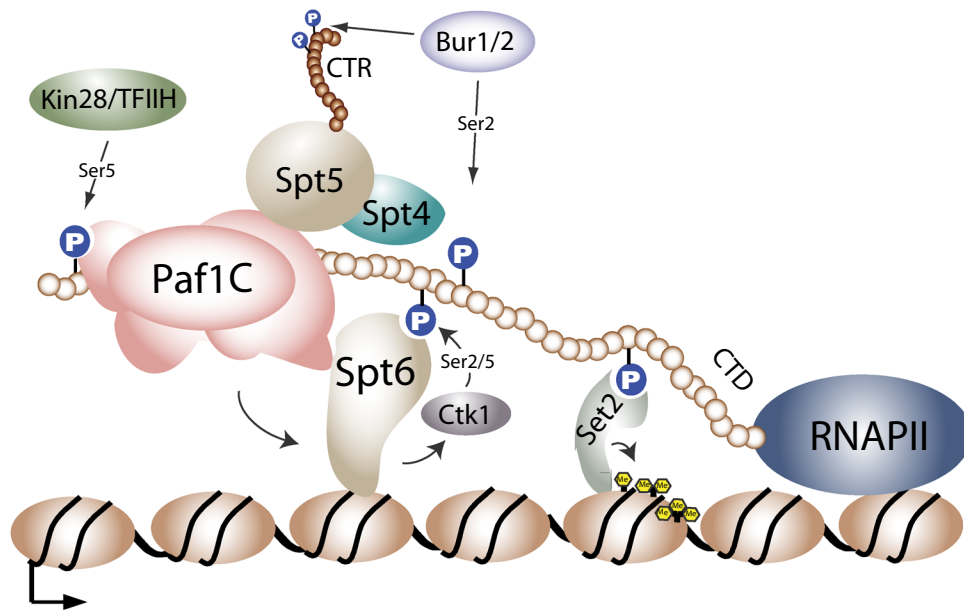
### The Paf1 complex

The Paf1 complex, which associates with RNAPII and facilitates transcription elongation, coordinates multiple histone PTMs during transcription. Specifically, the Paf1 complex is required for H3K36me3 [40, 41] in addition to H2BK123 mono-ubiquitylation (H2BK123ub1) [57, 58]—a chromatin mark that is necessary for H3K4 and H3K79 methylation. In the case of H2BK123ub1, the histone modification domain of Rtf1 (a member of the Paf1 complex that links Paf1 to RNAPII; [59]) directly associates with Rad6/Bre1, thus providing a molecular basis for how Paf1 regulates H2Bub1 [60]. However, in the case of Set2 methylation, the role of Paf1 is likely indirect. Evidence shows that a loss of Paf1 complex members leads to loss of Ser2 CTD phosphorylation [45, 61]. In addition, Spt6 is likely stabilized via direct interaction with Paf1 [45], which co-stabilizes Ctk1. Because a loss of Paf1 complex function leads to a reduction in Spt6 and Ctk1 levels, the levels of Set2, which is dependent on Ser2 CTD

phosphorylation, are also affected [45, 46]. Together, these data suggest a model wherein the Paf1 complex helps to stabilize and/or recruit Spt6 and Ctk1 to chromatin. Once recruited to chromatin, Ctk1 can then phosphorylate the RNAPII CTD for binding by Set2 (Fig. 2).

### Histone H3K36 demethylation

While there is a great deal of regulation surrounding the ability of Set2 to actively methylate histones, H3K36me is also dynamic and regulated by the actions of histone demethylases. In budding yeast, there are two predominant H3K36me demethylase enzymes, Jhd1, which acts on H3K36me1/2, and Rph1, which acts on H3K36me2/3 [62–65]. While these proteins both show demethylase activity *in vivo*, the role of active demethylation is still poorly understood. In humans, KDM2A has been shown to remove H3K36me2/3. Intriguingly, this protein contains a ZF-CXXC motif that serves to target it to nonmethylated CpG dinucleotides in linker regions and CpG islands, thereby removing H3K36me2/3 in promoter regions [66, 67]. In addition, KDM2A also has been shown to interact with HP1 through its chromoshadow domain [68], thus allowing KDM2A to be targeted to H3K9me3-marked



**Fig. 2** The regulation of Set2 and H3K36 methylation. Many factors come together to regulate Set2 and the production of H3K36 methylation in budding yeast. Kin28 of TFIIF initiates transcriptional initiation through the phosphorylation of the RNA polymerase II (RNAPII) C-terminal domain (CTD) at serine 5 (Ser5). This phosphorylation is important for the recruitment of Spt4/5, the Bur1 kinase, and the Paf1 complex (Paf1C) that promote transcription elongation and downstream histone modifications such as H2BK123 mono-ubiquitylation and H3K4 and K79 methylation. In addition to Kin28, Bur1 also functions to recruit the Paf1 complex, most likely

through the combined actions of Bur1 activity on the CTD and the C-terminal repeat region (CTR) in Spt5. Significantly, Paf1C associates with Spt6 and stabilizes Spt6 protein levels, which is required for stabilization and function of the major Ser2 CTD kinase, Ctk1—thus creating a feed forward circuit as Ser2 CTD phosphorylation is also required for binding and stability of Spt6. These actions drive robust CTD Ser2 phosphorylation that is required for SRI-CTD interaction, which stabilizes Set2 for nucleosomal H3K36 methylation (see text for details)

nucleosomes and mediate chromatin silencing. Thus, it will be important in the future to define how these two activities are coordinated and work together in chromatin. Another human protein involved in H3K36me<sub>2/3</sub> removal is KDM4A [69]. This enzyme, which also removes H3K9me<sub>3</sub>, is highly amplified in human tumors [70] and is localized in the nucleolus where it regulates rDNA transcription [71]. Future studies will be needed to determine which functions of KDM4A are specifically mediated through its action on H3K36.

### HOTAIR and miR-106b-5p

SETD2 is also regulated at the transcriptional and post-transcriptional levels. The long noncoding RNA HOTAIR can bind to the promoter region of SETD2 and inhibit the recruitment of pro-transcription factors such as CREB, P300, and RNAPII [72]. At the post-transcriptional level, another non-coding RNA, miR-106b-5p, binds SETD2 mRNA and inhibits its translation in clear cell renal cell carcinoma (ccRCC) [73]. As discussed below, SETD2 is frequently mutated and/or inactivated in ccRCCs and other cancer types. Thus, miR-106b-5p and HOTAIR, which themselves play roles in cancer development [74, 75], may do so also, in part, through SETD2 inactivation and loss of H3K36me<sub>3</sub>. Further work is needed to identify other RNA factors that regulate SETD2 levels, and whether these RNA-regulating pathways function in normal aspects of cellular biology to regulate SETD2 and H3K36me<sub>3</sub>.

## Effector proteins that interact with unmodified or methylated H3K36

### Histone chaperones and Asf1

During transcription, nucleosomes are disassembled but then replaced to maintain proper chromatin structure. A number of chaperones contribute to nucleosome disassembly and reassembly during transcription, including Spt6, Spt16 of the FACT complex, and Asf1 [48, 76–78]. Asf1, in particular, is a histone chaperone for H3 and H4 that deposits newly synthesized histones modified with H3K56 acetylation (H3K56ac) during transcription and replication [76, 79, 80]. Accordingly, H3K56ac is a marker for histone deposition during replication and histone exchange during transcription elongation. Interestingly, cells deleted for *ASF1* also show a depletion of H3K36me in some strains [81], indicating a connection between histone eviction/deposition and Set2. Further, cells deleted for *SET2* (*set2Δ*) show increased H3K56ac at the 3'-ends of genes in G1-arrested cells [82], indicating increased levels of histone exchange. Conversely, cells lacking Asf1 exhibited

slightly lower levels of histone acetylation over gene bodies, indicating that Asf1 and Set2 have opposing functions, and a combined loss of Asf1 and Set2 results in a significant decrease in histone acetylation [82]. Critically, Asf1, and to a lesser extent Spt6 and Spt16, show decreased affinity for peptides that are marked with H3K36me [82], suggesting a model wherein Set2-mediated H3K36me may actively inhibit histone exchange after transcription by preventing Asf1 (and other histone chaperones) from binding to and removing H3K36me histones (Fig. 3a). Thus, H3K36 methylation may serve to physically aid in preventing histone exchange. Further work is needed to understand precisely how H3K36me-marked nucleosomes are refractory to histone chaperones and whether active H3K36 demethylation is needed to allow for nucleosome disruption.

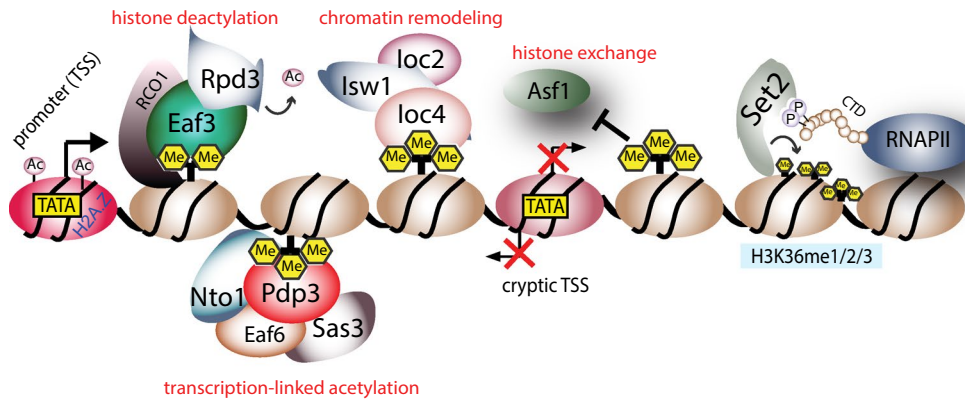
### Ioc4 and the Isw1b chromatin-remodeling complex

Yeast Ioc4 is a member of the Isw1b ATP-dependent chromatin-remodeling complex [83]. Isw1 is the catalytic subunit of the complex, and it forms two distinct complexes *in vivo*, one with Ioc4 (Isw1b) and one without (Isw1a). Ioc4 has a PWWP domain, named for conserved proline and tryptophan residues that, in humans, preferentially binds to H3K36me<sub>3</sub> [84]. The PWWP domain of Ioc4 also binds H3K36me<sub>3</sub> [85, 86]. This interaction facilitates the recruitment of Isw1b to the 3' regions of gene bodies, where Isw1 can then slide nucleosomes along the DNA [86]. In the absence of Set2 or H3K36me, Isw1b can no longer associate properly with chromatin [85, 86], creating regions of improperly placed nucleosomes that are poor substrates for the Rpd3 histone deacetylase complex (discussed below) [87]. Although the mechanism is not fully understood, Isw1b also functions, in concert with Set2, to prevent histone exchange and limit H4 acetylation [86], again stressing the importance of Set2 and H3K36me in limiting both of these actions once RNAPII has successfully transcribed a gene.

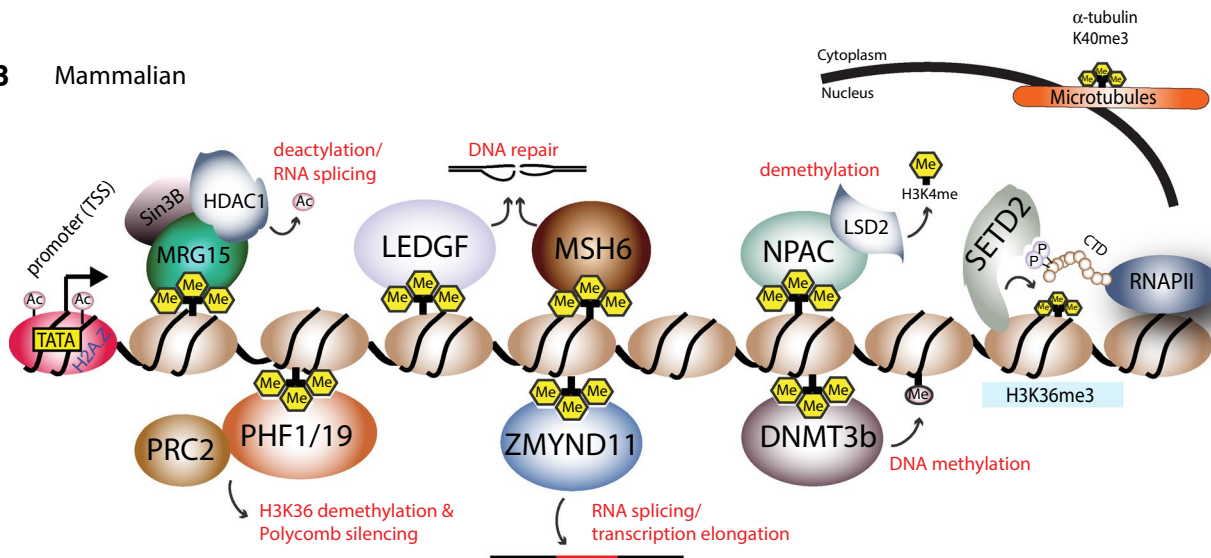
### Pdp3 and the NuA3b complex

In addition to *IOC4*, *PDP3* is a second yeast gene containing a PWWP domain. Pdp3 is a member of the NuA3b histone acetyltransferase (HAT) complex [88]. Like Ioc4, Pdp3 binds preferentially to H3K36me<sub>3</sub> and is dependent upon this mark to associate with chromatin. In the absence of Pdp3, genes that are targeted by NuA3 [89] experience marked reductions in the levels of their transcription, indicating that NuA3b plays an important role in transcription. It is not known how this complex is regulated by H3K36me<sub>3</sub> binding or whether it acetylates histones (perhaps H3K14 as does NuA3a) or a non-histone substrate. In addition, how NuA3b functions with Rpd3S or

**A Budding Yeast**



**B Mammalian**



**Fig. 3** H3K36me-interacting proteins in budding yeast and humans. **a** In *S. cerevisiae*, Set2 carries out mono-, di-, and tri-methylation of H3K36. Several functions have been ascribed to H3K36me in yeast, which include prevention of inappropriate or “cryptic” transcription, DNA damage repair and mRNA splicing. How H3K36me contributes to these processes is not fully understood, but a variety of effector proteins and complexes that bind to, or are repelled from H3K36me have been identified. These include the Rpd3S complex that binds H3K36me2/3 via the chromodomain of Eaf3 to deacetylate histones, the Isw1b chromatin-remodeling complex that binds H3K36me3 via the PWWP domain of Loc4 to promote proper nucleosome spacing across the gene body, the NuA3 complex that binds H3K36me3 via the PWWP domain of Pdp3 to promote acetylation and transcription of NuA3-specific target genes, and Asf1 (and other chaperones not shown for simplicity), which is a histone chaperone repelled by H3K36me3, thus preventing histone exchange and ensuring that nucleosomes are stable once incorporated into chromatin. Studies suggest that Rpd3S, Isw1b, and histone chaperones such as

Asf1 function together to regulate nucleosome deposition, nucleosome positioning, and prevent histone exchange in the wake of RNA polymerase II transcription. In the absence of Set2, nucleosomes are not positioned correctly, histone exchange and instability is increased, and normally obscured promoters become active. **b** In human cells, SETD2 is the sole enzyme for H3K36me3. In addition to the recruitment of a homologous Rpd3S complex containing MRG15, H3K36me3 recruits a variety of additional effector proteins and complexes that mediates transcription elongation, splicing, DNA methylation and histone deacetylation. These effectors include LEDGF, ZMYND11, DNMT3b, NPAC/LSD2, MSH6, and PHF1/19 (see Figure and main text for details on their functions). Recently, SETD2 was shown to methylate  $\alpha$ -tubulin, the first non-histone substrate identified for any H3K36 methylase. Trimethylation of  $\alpha$ -tubulin was shown to be important for proper mitosis and genome integrity, providing a potential explanation for how SETD2 functions as a tumor repressor

Isw1b is also unknown, and resolution of the opposing role H3K36me plays in both histone acetylation and deacetylation will be intriguing.

**Eaf3 and the Rpd3S complex**

Rpd3S, a histone deacetylase complex (HDAC), contains two proteins with reader domains that interact with chromatin: Eaf3 and Rco1. Eaf3 contains a chromo

domain that recognizes H3K36me<sub>2/3</sub> [90–94], and Rco1 contains a dual plant homeodomain (PHD) finger motif that is necessary for chromatin engagement [95] and a homo-dimerization domain [96]. Both PHD fingers are essential for chromatin binding and Rpd3S function, and both domains prefer to bind the unmodified N-terminus of H3 [97]. Interestingly, Rco1 exists as a dimer within the Rpd3S complex, adding two additional chromatin contacts to the complex [96]. Like the Isw1b complex, Eaf3's recognition of H3K36me helps to localize the Rpd3S complex to the 3'-ends of gene bodies [90, 94, 95] where it can deacetylate histone tails. In the absence of either Set2, Rco1, or Eaf3, H4 acetylation increases at the 3'-ends of gene bodies [90, 91, 94, 95] leading to inappropriate transcription initiation at cryptic promoters within the gene bodies.

Although Eaf3 and Rco1 play an important role in localizing Rpd3S across the genome, Rpd3S can be recruited to gene bodies independently of H3K36me binding. Like Set2, Rpd3S can be recruited to chromatin via the phosphorylated CTD of RNAPII, specifically the Ser2/Ser5 dually phosphorylated form [98, 99]. However, without Eaf3 binding to H3K36me<sub>2/3</sub>, or without the PHD fingers of Rco1, the Rpd3S complex is catalytically inactive [95, 97]. Recent work has suggested that H3K36me stimulates a conformational change in Rpd3S that increases its affinity for chromatin, perhaps stimulating its enzymatic activity [100].

Interestingly, Rpd3S prefers an H3K36me di-nucleosome substrate [101]. Rpd3S deacetylase activity is further promoted when these nucleosomes are properly spaced from one another [87]. The correct spacing is likely ensured by the nucleosome sliding activity of Isw1b, which could explain the increased histone acetylation levels in *isw1Δ* and *ioc4Δ* cells [86]. Without proper spacing, and even with proper H3K36me, the affinity of Rpd3S for chromatin is decreased to a degree where the naturally low binding affinity of Eaf3 for H3K36me<sub>2/3</sub> is not able to sufficiently engage nucleosomes and stimulate Rpd3S's deacetylase activity. Further, Rpd3S also depends on the PHD fingers in Rco1 to maintain its localization on chromatin [95, 97]. If any of these conditions are not met, the net result is increased histone acetylation in gene bodies. Collectively, these findings indicate that at least one function of H3K36me is to maintain nucleosome stability by repelling histone chaperones like Asf1 and ensure low levels of histone acetylation by recruiting the Isw1b chromatin-remodeling complex and the Rpd3S HDAC. These activities prevent RNAPII from engaging cryptic promoters within gene bodies (Fig. 3a).

## Human H3K36me<sub>3</sub>-associated factors

In humans, H3K36 is methylated by a variety of enzymes [102–104] including SETD2, which similarly associates, like yeast Set2, with elongating RNAPII to tri-methylate H3K36 [17, 31, 105, 106]. Similar to yeast, H3K36me<sub>3</sub> recruits an Rpd3S-related HDAC complex that associates via the Eaf3 homolog MRG15 [107, 108]; MRG15 also maintains reduced histone acetylation levels and functions in alternative mRNA splicing [107, 109]. In addition, SETD2/H3K36me<sub>3</sub> recruits a number of other chromatin-associated proteins that share a PWWP domain and mediate a wide variety of cellular processes including transcription elongation, heterochromatin formation, mRNA splicing, and DNA repair. These other proteins include ZMYND11 that regulates pre-mRNA splicing and transcription elongation [110, 111], the DNA methyltransferase DNMT3b that regulates gene body DNA methylation [112, 113], PHF1/PHF19 of the polycomb repressive complex 2 (PRC2) complex that recruits H3K27 methylation to silence developmental genes marked for repression [114–116], NPAC, which associates with a H3K4 demethylase [84, 117], LEDGF, a DNA repair factor that functions in homologous recombination [118], and MSH6, a factor required for mismatch repair [119] (Fig. 3b). The functions of these human and yeast proteins are discussed further below in the context of the biological processes they operate in.

## Cryptic transcription

One of the best-understood functions of Set2 and H3K36me in yeast is prevention of cryptic transcription across the genome. A cryptic transcript is traditionally defined as a transcript that originates from inside a gene body as opposed to originating from the canonical 5' promoter region [120]. Cryptic transcripts can arise from both the sense and antisense direction [121, 122], and they seem to occur preferentially from longer genes that are weakly transcribed [123]. Cryptic transcripts can be further classified based on their stabilities and how they are degraded. Generally, cryptic transcripts are quickly degraded by the exosome [124] and are referred to as cryptic unstable transcripts (CUTs) [121, 122]. These highly unstable transcripts can be observed and studied only when the exosome is compromised [124]. One such class of CUTs is the Xrn1-sensitive unstable transcripts that are reliant on the Xrn1 5'–3' exonuclease for degradation [125]. Additionally, a new class of cryptic transcripts, Set2-repressed antisense transcripts, arises selectively in *set2Δ* cells [126]. Finally, the stable unannotated transcripts (SUTs) comprise a separate class of transcripts that are much more stable [122]. Genome-wide studies in yeast have shown that



cryptic transcripts can arise from within the 3' region of gene bodies, but result at an even higher frequency from bi-directional transcription events at promoters [121, 122, 127]. Excitingly, cryptic transcription has recently been described in mammalian cells, and was shown to be linked to the recruitment and function of DNMT3b via H3K36me3 [128]. In both organisms, it remains to be determined what function, if any, these transcripts may serve.

There are two main mechanisms by which cryptic transcripts are thought to arise: increased histone acetylation and decreased numbers or mis-localization of histones across the gene body. As discussed above and further below, Set2 and H3K36me play an important role in preventing both of these aberrant phenomena.

Isw1b and Rpd3S play pivotal roles in ensuring that gene bodies remain hypo-acetylated after passage of the elongating RNAPII. A deletion of *ISWI* or *IOC4* results in increased histone acetylation and increased cryptic transcription [86]. This occurrence is likely due to the fact that nucleosomes are no longer positioned correctly for Rpd3S to engage a di-nucleosome pair [87]. Biochemical experiments have demonstrated that Rpd3S has a preferred linker length of ~50 base pairs of DNA [87]. The addition of Isw1 leads to increased Rpd3S deacetylase activity [87]. Isw1b and other chromatin remodelers are likely necessary to accurately prime the chromatin template for deacetylation by Rpd3S. In total, H3K36me serves to both activate Rpd3S activity through the binding of Eaf3's chromo-domain and, through Isw1b recruitment via the PWWP domain of Ioc4, ensure that nucleosome spacing is ideal for Rpd3S binding. These functions maintain a repressive chromatin environment behind RNAPII, and they reinforce transcription in the sense direction [127].

The other mechanism by which cryptic transcription is repressed is by ensuring that nucleosomes are properly restored after the elongating RNAPII complex has moved through the gene body. Spt6, Asf1, and FACT are all important in this process. These chaperones prevent cryptic transcription by reassembling nucleosomes behind the elongating RNAPII complex [120]. If the functions of these chaperones are compromised, there are fewer intact nucleosomes on chromatin, allowing RNAPII to access cryptic promoters [54, 55]. Further, nucleosomes marked with Set2 methylation are refractory to Asf1, Spt6, and FACT binding, thus ensuring that, once nucleosomes are assembled into chromatin behind RNAPII, they remain assembled [82].

Increased histone acetylation and defects in nucleosome reassembly are likely not the only two mechanisms that give rise to cryptic transcripts. Deletions of many chromatin modifying, remodeling, and maintenance factors promote cryptic transcription [55, 129]. Interestingly, the types of cryptic transcripts produced from

these different deletion mutants are different [86], offering further support for different underlying mechanisms regulating cryptic transcript production. Curiously, widespread production of cryptic transcripts does not necessarily have a deleterious effect on cell growth. *set2Δ* cells are viable, as are many other deletions that lead to the production of cryptic transcripts. Further, under normal laboratory conditions, a deletion of *SET2* or other chromatin factors results in relatively few changes to the transcriptome, i.e., only ~80 genes are up- or down-regulated in a *SET2* deletion [130]; thus, production of cryptic transcripts does not drive large-scale transcriptional change. However, there is limited evidence that severe nutrient deprivation can produce cryptic transcripts in wild-type cells [55]. This finding suggests that cryptic transcripts could act as a defense mechanism or have relevant functions during cellular stress.

Although, historically, cryptic transcripts have not been ascribed functions, there are currently two thoughts as to their possible cellular effects or roles. First, cryptic transcripts could inhibit transcription. Genes containing antisense transcripts that overlapped with their promoters exhibited statistically significant drops in full-length transcript amounts [131], which in some cases altered protein levels. Consistent with a transcriptional inhibition function, recent studies comparing the activation of genes upon carbon source shift showed a potential for antisense transcripts to cross sense promoters and deposit Set2-mediated H3K36me that, in turn, reduced the transcription of the sense genes [132]. Additional studies, under a variety of growth conditions, are needed to ascertain whether this phenomenon is a normal consequence of cryptic transcription or whether it is limited to nutrient shift. It is also necessary to determine whether the production of antisense transcripts during nutrient shift is required for the proper cellular response to stress.

The other function of cryptic transcripts could be the production of cryptic proteins. Using Spt6 mutation that produces a very strong cryptic transcription phenotype, the Winston lab observed the translation of a select few cryptic transcripts [55]. Although no functions have been assigned to these proteins, they raise the intriguing possibility that the increasing complexity that we are just beginning to uncover in the transcriptome could extend to the proteome. One particularly compelling hypothesis is that cryptic proteins could behave as dominant negative variants of their wild-type counterparts. This conjecture is primarily relevant to in-frame, sense, cryptic transcripts, as they would theoretically encode fully functional C-terminal domains of proteins, yet lack the N-terminal domains. Quantitative mass spectrometry data sets examining the proteomes of wild-type and cryptic transcript-producing cells will be needed to understand how wide-spread the production of

cryptic proteins is, and to ascertain what cellular role(s) they may play.

### Cryptic transcription and aging

In addition to the repression of cryptic transcription, Set2 and H3K36me have other critical functions. One important function is regulation of aging, likely through the maintenance of transcriptional fidelity across the genome [133]. Specifically, the Berger group showed that a loss of Set2 or a mutation of H3K36 decreased the lifespan of yeast cells and, critically, removal of Rph1, an H3K36me3 demethylase, significantly extended the lifespan of wild-type cells. Interestingly, as yeast cells aged, they lost H3K36me genome-wide, and a significant increase in cryptic transcription was observed. The loss in H3K36me as yeast cells age could be due to a global loss in histones and/or an increase in H4K16 acetylation. Older yeast cells lose as much as 50% of their histones as they age, and they gain H4K16ac, further increasing the likelihood of transcriptional de-repression genome-wide [134]. Critically, these aforementioned functions are likely conserved in higher eukaryotes, as the *C. elegans* Set2 homolog MET-1 also extends lifespan [133, 135, 136]. In summary, in yeast and worms, H3K36me regulates the transcriptome by preventing cryptic transcription, thus decreasing transcriptional noise. This regulation ensures that genes are properly expressed throughout the organism's lifespan, thereby promoting longevity.

### Other functions of Set2/H3K36me and their evolutionary conservation

#### RNA splicing

A role of H3K36me in regulating splicing was first observed in human cells [109]. H3K36me3 regulates exon choice in the *FGFR2* gene. In epithelial cells, exon IIIb is included in the mRNA transcript, whereas in mesenchymal cells, exon IIIc is included. Overexpression or siRNA-mediated knockdown of *SETD2* demonstrated a shift in exon inclusion in both cell types. The mechanism of exon choice is recruitment of the splicing factor PTB via an interaction with MRG15, the human *EAF3* homolog [109]. MRG15 is recruited to *FGFR2* in an H3K36me3-dependent manner, and subsequently recruits PTB, which guides exon choice.

A genome-wide role of SETD2 in splicing has been demonstrated by next generation sequencing techniques [111, 137]. Without SETD2, a plethora of RNA splicing defects was observed, including intron retention and aberrantly spliced genes. Interestingly, nucleosome positioning

was altered at many sites of RNA processing defects, suggesting that H3K36me's chromatin remodeling function and repression of histone exchange may also be conserved in human cells.

ZMYND11 is another factor that binds H3K36me3 and regulates splicing and transcription elongation [110, 111]. This protein contains three tandemly arranged chromatin reader domains: a PHD, a bromodomain, and a PWWP domain. The PWWP domain selectively binds to H3K36me3, and, in concert with the adjacent bromodomain, recognizes residue S31 in the transcription-associated H3.3 variant, suggesting a specific role in transcription. In fact, ZMYND11 associates with RNA splicing regulators and RNA-seq data from a ZMYND11 knock-down showed increased intron-retention, further arguing for a role in mRNA splicing. Interestingly, ZMYND11 is down regulated in many cancers and, when overexpressed, ZMYND11 suppressed cancer cell growth [110].

Finally, similar to mammalian cells, Set2 also regulates splicing in yeast. Mutants lacking H3K36me exhibited reduced splicing efficiency across the genome [138]. Interestingly, correct splicing was dependent on H3K36me2 and association of Set2 with the CTD of RNAPII. The necessity of the SRI domain for proper splicing suggests that Set2 may act co-transcriptionally to recruit splicing factors, and the requirement of H3K36me2 suggests that Rpd3S may also play an important role in splicing.

#### DNA damage response

As DNA is packaged into chromatin, histones play a key role in the DNA damage response (DDR) pathways. Intriguingly, Set2 and H3K36me are necessary for DDR, although the mechanisms differ between organisms.

In fission and budding yeast, cells lacking Set2 and H3K36me display strong sensitivity to DNA damaging agents [139–141]. In both systems, Set2 regulates chromatin compaction and limits resection, i.e., the creation of single-stranded DNA at sites of DNA damage that allows binding of DDR proteins such as RPA. In the absence of H3K36me, the H3K36 residue is instead available for acetylation by Gcn5, resulting in chromatin decompaction. Further, limiting resection promotes non-homologous end joining (NHEJ) [139, 140]. In budding yeast, there were also defects in DDR signaling pathways in *set2Δ* cells [139], suggesting that the chromatin state surrounding the damage site is critical for recruiting DDR proteins and permitting their proper signaling activities. In fission yeast, H3K36me is cell cycle regulated, further delineating the choice between homologous recombination (HR) and NHEJ [140].

In human cells, HR is defective in SETD2 mutant cells [142–144]. This defect appears to be the result of a lack of RPA and RAD51 binding to damage sites [143, 144].

Further, upon loss of H3K36me<sub>3</sub>, DNA damage persists in cells, likely due to an inability to activate downstream signaling proteins, such as p53, that are critical for efficient repair [142].

In human cells, H3K36me also regulates DNA mismatch repair (MMR) by recruiting MSH6 [119, 143]. MSH6, like many H3K36me<sub>3</sub> effector proteins, contains a PWWP domain that is essential for H3K36me<sub>3</sub> binding. Critically, cells deficient for SETD2 display microsatellite instability and increased mutation frequency, phenotypes typical of defects in DNA MMR. Thus, Set2/SETD2 and H3K36me play important roles in maintaining genome integrity and stability.

### Polycomb silencing

H3K36me regulates Polycomb-mediated gene silencing during mammalian development by recruiting Polycomb repressive complex 2 (PRC2) components PHF19 and PHF1 [114–116, 145]. PHF19 and PHF1 both have Tudor domains that recognize H3K36me<sub>3</sub> and bring PRC2 to actively transcribed and developmentally regulated regions of the genome, which then allows for H3K27me<sub>3</sub> to spread across the genes to induce transcriptional repression. Additionally, PHF19 associates with NO66, an H3K36me<sub>3</sub> demethylase that further reinforces the silencing of these loci [116].

### DNA methylation

H3K36me is necessary for establishing DNA methylation in gene bodies. DNA methylation is an essential regulator of gene expression; hyper-methylation of gene promoters results in repression and is a hallmark of cancer. Both of the de novo DNA methyltransferases, DNMT3a and DNMT3b, contain PWWP domains that bind to H3K36me<sub>3</sub> [146, 147] and, further, are necessary for their DNA methyltransferase activities [112, 146]. In the absence of SETD2, DNA methylation is not properly targeted to transcribed regions in the genome, which has been attributed primarily to the loss of DNMT3b targeting by H3K36me<sub>3</sub> [112, 113, 148]. In the context of transcribed regions, the function of DNA methylation is poorly understood, but studies show that gene body DNA methylation is likely important for transcription elongation and splicing [149]. Thus, the ability of SETD2 and H3K36me<sub>3</sub> to promote transcription elongation and splicing likely depends, in part, on maintaining DNA methylation. In addition, and as mentioned above, DNA methylation also prevents spurious transcription initiation within the gene body in mammalian cells [128]. Future work is needed to further understand the function of DNA methylation in transcribed regions.

### Cell division via tubulin methylation by SETD2 at K40

In addition to its chromatin-based functions, SETD2 maintains genomic stability by ensuring the proper segregation of cells during mitosis [150]. During mitosis, SETD2 methylates tubulin at K40, and ablation of SETD2 leads to a plethora of mitotic defects. Most likely, some of the oncogenic phenotypes observed in SETD2-mutated cells are due to the genomic instability caused by tubulin defects during mitosis. Also, additional non-histone targets for SETD2 are likely to be discovered because many other histone methyltransferases (e.g., G9a) have multiple substrates [151].

### Animal development

Consistent with SETD2 and H3K36me having many cellular functions, multiple animal models have shown that SETD2 and the H3K36 histone residue are necessary for animal development. The requirement for SETD2 and H3K36me was first shown for *Drosophila*; RNAi knock-down of dSet2 caused lethality during the larval stage of development [14, 152]. Additionally, when all of the canonical H3K36 residues in the genome were mutated to arginine, flies were unable to complete development [153]. Thus, H3K36me is specifically critical for fly development.

SETD2 and H3K36me are also crucial for mammalian development. SETD2<sup>-/-</sup> knock-out mice die as embryos at E10.5–E11.5 due to defects in angiogenesis [154]. These embryos display vascular defects in the embryo, the yolk sac, and the placenta. RNAi knock-down experiments in murine embryonic stem cells also reveal defects in differentiation due to misregulation of the Fgfr3/Erk pathway [155].

These results demonstrate that the enzyme that mediates H3K36 methylation and the residue that accepts this methylation are critical for animal development; however, the precise mechanisms underlying their importance remain to be elucidated. Further, it is likely that the failure to methylate non-histone targets could contribute to the lethality observed in these animals. These exciting questions await experimental answers.

### The role of SETD2 and H3K36 methylation in cancer

SETD2 is mutated in up to 15% of patients with clear cell renal cell carcinoma (ccRCC) [156–158]. SETD2 is located on chromosome 3p—a region commonly deleted in ccRCC tumors [158]. ccRCC tumors are very heterogeneous, but mutations in SETD2 arise frequently and independently in a single tumor [157]. SETD2 mutations have also been observed in acute leukemias, in bladder cancer, and in

glioblastoma [159–161], indicating a broad need for maintaining SETD2 to prevent tumorigenesis.

Although, it is not known, biochemically, exactly how SETD2 functions as a tumor suppressor, the abilities of SETD2 and H3K36me<sub>3</sub> to regulate splicing, DNA methylation, chromosome segregation, and DNA damage repair are likely candidates that underlie SETD2's role in tumor suppression. SETD2 plays a critical role in maintaining transcriptional fidelity. In the absence of SETD2, mRNA processing defects occur at as many as 25% of expressed genes across the genome [137]. Further, in ccRCC, there are many transcription termination defects in SETD2-mutant cancer cell lines [162]. These termination defects lead to the creation of chimeric transcripts, some involving oncogenes, providing yet another potential mechanism that could lead to cancer development. Together, the defects in mRNA processing, transcription termination, DNA methylation, and impaired DNA damage signaling likely provide an ideal environment for tumorigenesis.

Finally, it has recently been shown through exome sequencing that recurrent mutations at or near H3K36 can also lead to cancer [163–165]. In particular, H3K36 in the context of the H3.3 variant has been found mutated to methionine in chondroblastomas [165] and in other cancer types including head and neck squamous cell carcinoma and colorectal cancer [166]. Mutations at H3.3K36 result in genome-wide loss of H3K36 tri-methylation and can directly stimulate tumor formation [167, 168]. Similar to the mechanism by which K→M mutations at H3K9 and H3K27 “trap” or “poison” their respective methyltransferase enzymes [169–172], the H3.3K36M mutation rearranges the active center of SETD2 to provide a high affinity binding site for the H3.3K36M histone tail, thus explaining how a single histone mutation can impart a tumor phenotype [173]. As H3K36me is lost from the genome, a redistribution of H3K27me<sub>3</sub> by PRC2 occurs across the genome [168]. Significantly, H3K27me<sub>3</sub> redistribution dilutes the polycomb repressive complex 1 (PRC1) that associates with H3K27me<sub>3</sub> across the genome and results in increased gene expression [168]. This redistribution disrupts the normal differentiation process by locking cells in an undifferentiated state.

In addition to mutations at H3.3K36, mutations at adjacent residues (i.e., G34-to-R/V and G34-to-W/L) have also been reported in a variety of cancers including pediatric non-brain stem gliomas and tumors of the bone [163–165]. Unlike the H3.3K36M mutation, however, G34W/L and G34R/V mutations do not abolish global levels of H3K36me<sub>3</sub> [170, 174]. Rather, they impact the nucleosomes that contain this mutation and result in a redistribution of the H3K36me<sub>3</sub> mark across the genome that reprograms the transcriptome in ways that promote

tumorigenesis [170, 174]. Collectively, these findings show that mutating either SETD2 or H3K36 is sufficient to promote cancer progression.

## Conclusions

H3K36me is a histone PTM that is conserved from yeast to humans, suggesting that it is extremely important for proper cellular function. Although great strides have been made recently to elucidate the participation of H3K36me in various cellular processes, we are just beginning to understand its functions in these newly discovered areas. The recent insights into Set2/SETD2 biology have raised many questions that will be explored over the coming years. Significantly, we do not fully understand how all of the conserved domains in Set2/SETD2 (e.g., the WW and CC domains) contribute to Set2/SETD2 targeting to chromatin and to its enzymatic activity toward histone and non-histone substrates. We also do not fully understand the distinct roles of each H3K36 methylation state. Each methyl state likely recruits distinct effector proteins, but the majority of effector proteins discovered to date (i.e., those with PWWP domains) prefer to bind to H3K36me<sub>3</sub>. Thus, it is likely that we have not yet identified the full gamut of H3K36me effector proteins and their functions.

Another significant question is why cells invest considerable energy in repressing cryptic transcription? Set2 and a plethora of other chromatin factors function to prevent these transcripts from arising. So we ask: Are cryptic transcripts deleterious to genome stability or integrity? Are these transcripts translated? If yes, how do cryptic proteins impact cellular biology? The answers to these exciting questions may also shed light on the role of H3K36me in suppressing tumorigenesis. Along these lines, although SETD2 and H3.3K36 are mutated in multiple types of cancer, we do not fully understand H3.3K36's role as a tumor suppressor. In addition, the recent discovery of a non-histone target for SETD2 will undoubtedly spur the search for other non-histone SETD2 targets, which, along with H3K36me and tubulin methylation, may function to maintain genome stability. Further work will elucidate how these methylation events contribute to chromatin and cellular function.

**Acknowledgements** We thank Deepak Jha, Howard Fried, and Julia DiFiore for helpful comments and suggestions. The work in the Strahl lab is supported by NIH grants GM110058 and GM68088 to BDS.

**Authors' contributions** SLM and BDS conceived, drafted, and wrote the manuscript.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no competing interests.

## References

- Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389(6648):251–260
- Kornberg RD (1974) Chromatin structure: a repeating unit of histones and DNA. *Science* 184(4139):868–871
- Kornberg RD, Lorch Y (1999) Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* 98(3):285–294
- Oudet P, Gross-Bellard M, Chambon P (1975) Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* 4(4):281–300
- Bannister AJ, Kouzarides T (2011) Regulation of chromatin by histone modifications. *Cell Res* 21(3):381–395
- Rothbart SB, Strahl BD (2014) Interpreting the language of histone and DNA modifications. *Biochim Biophys Acta* 1839(8):627–643
- Ng MK, Cheung P (2016) A brief histone in time: understanding the combinatorial functions of histone PTMs in the nucleosome context. *Biochem Cell Biol* 94(1):33–42
- Black JC, Van Rechem C, Whetstine JR (2012) Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol Cell* 48(4):491–507
- Tessarz P, Kouzarides T (2014) Histone core modifications regulating nucleosome structure and dynamics. *Nat Rev Mol Cell Biol* 15(11):703–708
- Wozniak GG, Strahl BD (2014) Hitting the ‘mark’: interpreting lysine methylation in the context of active transcription. *Biochim Biophys Acta* 1839(12):1353–1361
- Weiner A, Hsieh TH, Appleboim A, Chen HV, Rahat A, Amit I, Rando OJ, Friedman N (2015) High-resolution chromatin dynamics during a yeast stress response. *Mol Cell* 58(2):371–386
- Ho JW, Jung YL, Liu T, Alver BH, Lee S, Ikegami K, Sohn KA, Minoda A, Tolstorukov MY, Appert A et al (2014) Comparative analysis of metazoan chromatin organization. *Nature* 512(7515):449–452
- Strahl BD, Grant PA, Briggs SD, Sun ZW, Bone JR, Caldwell JA, Mollah S, Cook RG, Shabanowitz J, Hunt DF et al (2002) Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol Cell Biol* 22(5):1298–1306
- Bell O, Wirbelauer C, Hild M, Scharf AN, Schwaiger M, MacAlpine DM, Zilbermann F, van Leeuwen F, Bell SP, Imhof A et al (2007) Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in *Drosophila*. *EMBO J* 26(24):4974–4984
- Wagner EJ, Carpenter PB (2012) Understanding the language of Lys36 methylation at histone H3. *Nat Rev Mol Cell Biol* 13(2):115–126
- Yi X, Jiang XJ, Li XY, Jiang DS (2015) Histone methyltransferases: novel targets for tumor and developmental defects. *Am J Transl Res* 7(11):2159–2175
- Edmunds JW, Mahadevan LC, Clayton AL (2008) Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J* 27(2):406–420
- Wang Y, Niu Y, Li B (2015) Balancing acts of SRI and an auto-inhibitory domain specify Set2 function at transcribed chromatin. *Nucleic Acids Res* 43(10):4881–4892
- Hacker KE, Fahey CC, Shinsky SA, Chiang YJ, DiFiore JV, Jha DK, Vo AH, Shavit JA, Davis IJ, Strahl BD et al (2016) Structure/function analysis of recurrent mutations in SETD2 reveals a critical and conserved role for a SET domain residue in maintaining protein stability and H3K36 trimethylation. *J Biol Chem* 291:21283–21295
- Du HN, Fingerman IM, Briggs SD (2008) Histone H3 K36 methylation is mediated by a trans-histone methylation pathway involving an interaction between Set2 and histone H4. *Genes Dev* 22(20):2786–2798
- Du HN, Briggs SD (2010) A nucleosome surface formed by histone H4, H2A, and H3 residues is needed for proper histone H3 Lys36 methylation, histone acetylation, and repression of cryptic transcription. *J Biol Chem* 285(15):11704–11713
- Sudol M, Chen HI, Bougeret C, Einbond A, Bork P (1995) Characterization of a novel protein-binding module—the WW domain. *FEBS Lett* 369(1):67–71
- Sudol M, Sliwa K, Russo T (2001) Functions of WW domains in the nucleus. *FEBS Lett* 490(3):190–195
- Xiao T, Hall H, Kizer KO, Shibata Y, Hall MC, Borchers CH, Strahl BD (2003) Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes Dev* 17(5):654–663
- Gao YG, Yang H, Zhao J, Jiang YJ, Hu HY (2014) Autoinhibitory structure of the WW domain of HYPB/SETD2 regulates its interaction with the proline-rich region of huntingtin. *Structure* 22(3):378–386
- Faber PW, Barnes GT, Srinidhi J, Chen J, Gusella JF, MacDonald ME (1998) Huntingtin interacts with a family of WW domain proteins. *Hum Mol Genet* 7(9):1463–1474
- Poon BP, Mekhail K (2011) Cohesin and related coiled-coil domain-containing complexes physically and functionally connect the dots across the genome. *Cell Cycle* 10(16):2669–2682
- Rebhed J, Revy P, Faure G, de Villartay JP, Callebaut I (2014) Expanding the SRI domain family: a common scaffold for binding the phosphorylated C-terminal domain of RNA polymerase II. *FEBS Lett* 588(23):4431–4437
- Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD (2005) A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol Cell Biol* 25(8):3305–3316
- Schaft D, Roguev A, Kotovic KM, Shevchenko A, Sarov M, Shevchenko A, Neugebauer KM, Stewart AF (2003) The histone 3 lysine 36 methyltransferase, SET2, is involved in transcriptional elongation. *Nucleic Acids Res* 31(10):2475–2482
- Li M, Phatnani HP, Guan Z, Sage H, Greenleaf AL, Zhou P (2005) Solution structure of the Set2-Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proc Natl Acad Sci USA* 102(49):17636–17641
- Vojnic E, Simon B, Strahl BD, Sattler M, Cramer P (2006) Structure and carboxyl-terminal domain (CTD) binding of the Set2 SRI domain that couples histone H3 Lys36 methylation to transcription. *J Biol Chem* 281(1):13–15
- Li B, Howe L, Anderson S, Yates JR 3rd, Workman JL (2003) The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* 278(11):8897–8903
- Youde ML, Kizer KO, Kisseleva-Romanova E, Fuchs SM, Duro E, Strahl BD, Mellor J (2008) Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36. *Mol Cell Biol* 28(16):4915–4926
- Suzuki S, Kato H, Suzuki Y, Chikashige Y, Hiraoka Y, Kimura H, Nagao K, Obuse C, Takahata S, Murakami Y

- (2016) Histone H3K36 trimethylation is essential for multiple silencing mechanisms in fission yeast. *Nucleic Acids Res* 44(9):4147–4162
36. Jones JC, Phatnani HP, Haystead TA, MacDonald JA, Alam SM, Greenleaf AL (2004) C-terminal repeat domain kinase I phosphorylates Ser2 and Ser5 of RNA polymerase II C-terminal domain repeats. *J Biol Chem* 279(24):24957–24964
  37. Suh H, Ficarro SB, Kang UB, Chun Y, Marto JA, Buratowski S (2016) Direct analysis of phosphorylation sites on the Rpb1 C-terminal domain of RNA polymerase II. *Mol Cell Biol* 36(2):297–304
  38. Qiu H, Hu C, Hinnebusch AG (2009) Phosphorylation of the Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters. *Mol Cell Biol* 29(6):752–762
  39. Liu Y, Warfield L, Zhang C, Luo J, Allen J, Lang WH, Ranish J, Shokat KM, Hahn S (2009) Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. *Mol Cell Biol* 29(17):4852–4863
  40. Chu Y, Sutton A, Sternglanz R, Prelich G (2006) The BUR1 cyclin-dependent protein kinase is required for the normal pattern of histone methylation by SET2. *Mol Cell Biol* 26(8):3029–3038
  41. Chu Y, Simic R, Warner MH, Arndt KM, Prelich G (2007) Regulation of histone modification and cryptic transcription by the Bur1 and Paf1 complexes. *EMBO J* 26(22):4646–4656
  42. Keogh MC, Podolny V, Buratowski S (2003) Bur1 kinase is required for efficient transcription elongation by RNA polymerase II. *Mol Cell Biol* 23(19):7005–7018
  43. Larabee RN, Krogan NJ, Xiao T, Shibata Y, Hughes TR, Greenblatt JF, Strahl BD (2005) BUR kinase selectively regulates H3 K4 trimethylation and H2B ubiquitylation through recruitment of the PAF elongation complex. *Curr Biol* 15(16):1487–1493
  44. Zhou K, Kuo WH, Fillingham J, Greenblatt JF (2009) Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. *Proc Natl Acad Sci USA* 106(17):6956–6961
  45. Dronamraju R, Strahl BD (2014) A feed forward circuit comprising Spt6, Ctk1 and PAF regulates Pol II CTD phosphorylation and transcription elongation. *Nucleic Acids Res* 42(2):870–881
  46. Fuchs SM, Kizer KO, Braberg H, Krogan NJ, Strahl BD (2012) RNA polymerase II carboxyl-terminal domain phosphorylation regulates protein stability of the Set2 methyltransferase and histone H3 di- and trimethylation at lysine 36. *J Biol Chem* 287(5):3249–3256
  47. Zhu K, Lei PJ, Ju LG, Wang X, Huang K, Yang B, Shao C, Zhu Y, Wei G, Fu XD et al (2016) SPOP-containing complex regulates SETD2 stability and H3K36me3-coupled alternative splicing. *Nucleic Acids Res* 45:92–105
  48. Bortvin A, Winston F (1996) Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science* 272(5267):1473–1476
  49. Clark-Adams CD, Winston F (1987) The SPT6 gene is essential for growth and is required for delta-mediated transcription in *Saccharomyces cerevisiae*. *Mol Cell Biol* 7(2):679–686
  50. Swanson MS, Carlson M, Winston F (1990) SPT6, an essential gene that affects transcription in *Saccharomyces cerevisiae*, encodes a nuclear protein with an extremely acidic amino terminus. *Mol Cell Biol* 10(9):4935–4941
  51. Diebold ML, Loeliger E, Koch M, Winston F, Cavarelli J, Romier C (2010) Noncanonical tandem SH2 enables interaction of elongation factor Spt6 with RNA polymerase II. *J Biol Chem* 285(49):38389–38398
  52. Sun M, Lariviere L, Dengl S, Mayer A, Cramer P (2010) A tandem SH2 domain in transcription elongation factor Spt6 binds the phosphorylated RNA polymerase II C-terminal repeat domain (CTD). *J Biol Chem* 285(53):41597–41603
  53. Yoh SM, Lucas JS, Jones KA (2008) The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. *Genes Dev* 22(24):3422–3434
  54. DeGennaro CM, Alver BH, Marguerat S, Stepanova E, Davis CP, Bahler J, Park PJ, Winston F (2013) Spt6 regulates intragenic and antisense transcription, nucleosome positioning, and histone modifications genome-wide in fission yeast. *Mol Cell Biol* 33(24):4779–4792
  55. Cheung V, Chua G, Batada NN, Landry CR, Michnick SW, Hughes TR, Winston F (2008) Chromatin- and transcription-related factors repress transcription from within coding regions throughout the *Saccharomyces cerevisiae* genome. *PLoS Biol* 6(11):e277
  56. Zhang L, Fletcher AG, Cheung V, Winston F, Stargell LA (2008) Spn1 regulates the recruitment of Spt6 and the Swi/Snf complex during transcriptional activation by RNA polymerase II. *Mol Cell Biol* 28(4):1393–1403
  57. Ng HH, Dole S, Struhl K (2003) The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B. *J Biol Chem* 278(36):33625–33628
  58. Wood A, Schneider J, Dover J, Johnston M, Shilatifard A (2003) The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. *J Biol Chem* 278(37):34739–34742
  59. Mueller CL, Porter SE, Hoffman MG, Jaehning JA (2004) The Paf1 complex has functions independent of actively transcribing RNA polymerase II. *Mol Cell Biol* 24(4):447–456
  60. Van Oss SB, Shirra MK, Bataille AR, Wier AD, Yen K, Vinayachandran V, Byeon IL, Cucinotta CE, Heroux A, Jeon J et al (2016) The histone modification domain of Paf1 complex subunit Rtf1 directly stimulates H2B ubiquitylation through an interaction with Rad6. *Mol Cell Biol* 36(4):815–825
  61. Nordick K, Hoffman MG, Betz JL, Jaehning JA (2008) Direct interactions between the Paf1 complex and a cleavage and polyadenylation factor are revealed by dissociation of Paf1 from RNA polymerase II. *Eukaryot Cell* 7(7):1158–1167
  62. Klose RJ, Gardner KE, Liang G, Erdjument-Bromage H, Tempst P, Zhang Y (2007) Demethylation of histone H3K36 and H3K9 by Rph1: a vestige of an H3K9 methylation system in *Saccharomyces cerevisiae*? *Mol Cell Biol* 27(11):3951–3961
  63. Fang J, Hogan GJ, Liang G, Lieb JD, Zhang Y (2007) The *Saccharomyces cerevisiae* histone demethylase Jhd1 fine-tunes the distribution of H3K36me2. *Mol Cell Biol* 27(13):5055–5065
  64. Tu S, Bulloch EM, Yang L, Ren C, Huang WC, Hsu PH, Chen CH, Liao CL, Yu HM, Lo WS et al (2007) Identification of histone demethylases in *Saccharomyces cerevisiae*. *J Biol Chem* 282(19):14262–14271
  65. Kim T, Buratowski S (2007) Two *Saccharomyces cerevisiae* JmjC domain proteins demethylate histone H3 Lys36 in transcribed regions to promote elongation. *J Biol Chem* 282(29):20827–20835
  66. Blackledge NP, Zhou JC, Tolstorukov MY, Farcas AM, Park PJ, Klose RJ (2010) CpG islands recruit a histone H3 lysine 36 demethylase. *Mol Cell Biol* 30(2):179–190
  67. Zhou JC, Blackledge NP, Farcas AM, Klose RJ (2012) Recognition of CpG island chromatin by KDM2A requires direct and specific interaction with linker DNA. *Mol Cell Biol* 32(2):479–489
  68. Borgel J, Tyl M, Schiller K, Pusztai Z, Dooley CM, Deng W, Wooding C, White RJ, Warnecke T, Leonhardt H et al (2016) KDM2A integrates DNA and histone modification signals through a CXXC/PHD module and direct interaction with HP1. *Nucleic Acids Res* 45:1114–1129

69. Whetstone JR, Nottke A, Lan F, Huarte M, Smolnikov S, Chen Z, Spooner E, Li E, Zhang G, Colaiacovo M et al (2006) Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* 125(3):467–481
70. Black JC, Manning AL, Van Rechem C, Kim J, Ladd B, Cho J, Pineda CM, Murphy N, Daniels DL, Montagna C et al (2013) KDM4A lysine demethylase induces site-specific copy gain and rereplication of regions amplified in tumors. *Cell* 154(3):541–555
71. Salifou K, Ray S, Verrier L, Aguirrebengoa M, Trouche D, Panov KI, Vandromme M (2016) The histone demethylase JMJD2A/KDM4A links ribosomal RNA transcription to nutrients and growth factors availability. *Nat Commun* 7:10174
72. Li H, An J, Wu M, Zheng Q, Gui X, Li T, Pu H, Lu D (2015) LncRNA HOTAIR promotes human liver cancer stem cell malignant growth through downregulation of SETD2. *Oncotarget* 6(29):27847–27864
73. Xiang W, He J, Huang C, Chen L, Tao D, Wu X, Wang M, Luo G, Xiao X, Zeng F et al (2015) miR-106b-5p targets tumor suppressor gene SETD2 to inactivate its function in clear cell renal cell carcinoma. *Oncotarget* 6(6):4066–4079
74. Tang L, Zhang W, Su B, Yu B (2013) Long noncoding RNA HOTAIR is associated with motility, invasion, and metastatic potential of metastatic melanoma. *Biomed Res Int* 2013:251098
75. Liu F, Gong J, Huang W, Wang Z, Wang M, Yang J, Wu C, Wu Z, Han B (2014) MicroRNA-106b-5p boosts glioma tumorigenesis by targeting multiple tumor suppressor genes. *Oncogene* 33(40):4813–4822
76. Schwabish MA, Struhl K (2006) Asf1 mediates histone eviction and deposition during elongation by RNA polymerase II. *Mol Cell* 22(3):415–422
77. Dengl S, Mayer A, Sun M, Cramer P (2009) Structure and in vivo requirement of the yeast Spt6 SH2 domain. *J Mol Biol* 389(1):211–225
78. Schwabish MA, Struhl K (2004) Evidence for eviction and rapid deposition of histones upon transcriptional elongation by RNA polymerase II. *Mol Cell Biol* 24(23):10111–10117
79. Adkins MW, Tyler JK (2004) The histone chaperone Asf1p mediates global chromatin disassembly in vivo. *J Biol Chem* 279(50):52069–52074
80. Rufiange A, Jacques PE, Bhat W, Robert F, Nourani A (2007) Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. *Mol Cell* 27(3):393–405
81. Lin LJ, Minard LV, Johnston GC, Singer RA, Schultz MC (2010) Asf1 can promote trimethylation of H3 K36 by Set2. *Mol Cell Biol* 30(5):1116–1129
82. Venkatesh S, Smolle M, Li H, Gogol MM, Saint M, Kumar S, Natarajan K, Workman JL (2012) Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. *Nature* 489(7416):452–455
83. Vary JC Jr, Gangaraju VK, Qin J, Landel CC, Kooperberg C, Bartholomew B, Tsukiyama T (2003) Yeast Isw1p forms two separable complexes in vivo. *Mol Cell Biol* 23(1):80–91
84. Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG et al (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142(6):967–980
85. Maltby VE, Martin BJ, Schulze JM, Johnson I, Hentrich T, Sharma A, Kobor MS, Howe L (2012) Histone H3 lysine 36 methylation targets the Isw1b remodeling complex to chromatin. *Mol Cell Biol* 32(17):3479–3485
86. Smolle M, Venkatesh S, Gogol MM, Li H, Zhang Y, Florens L, Washburn MP, Workman JL (2012) Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange. *Nat Struct Mol Biol* 19(9):884–892
87. Lee CH, Wu J, Li B (2013) Chromatin remodelers fine-tune H3K36me-directed deacetylation of neighbor nucleosomes by Rpd3S. *Mol Cell* 52(2):255–263
88. Gilbert TM, McDaniel SL, Byrum SD, Cades JA, Dancy BC, Wade H, Tackett AJ, Strahl BD, Taverna SD (2014) A PWWP domain-containing protein targets the NuA3 acetyltransferase complex via histone H3 lysine 36 trimethylation to coordinate transcriptional elongation at coding regions. *Mol Cell Proteomics* 13(11):2883–2895
89. Taverna SD, Ilin S, Rogers RS, Tanny JC, Lavender H, Li H, Baker L, Boyle J, Blair LP, Chait BT 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(5):785–796
90. Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia WJ, Anderson S, Yates J, Washburn MP et al (2005) Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* 123(4):581–592
91. Joshi AA, Struhl K (2005) Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol Cell* 20(6):971–978
92. Sun B, Hong J, Zhang P, Dong X, Shen X, Lin D, Ding J (2008) Molecular basis of the interaction of *Saccharomyces cerevisiae* Eaf3 chromo domain with methylated H3K36. *J Biol Chem* 283(52):36504–36512
93. Xu C, Cui G, Botuyan MV, Mer G (2008) Structural basis for the recognition of methylated histone H3K36 by the Eaf3 subunit of histone deacetylase complex Rpd3S. *Structure* 16(11):1740–1750
94. Keogh MC, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, Schuldiner M, Chin K, Punna T, Thompson NJ et al (2005) Cotranscriptional Set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* 123(4):593–605
95. Li B, Gogol M, Carey M, Lee D, Seidel C, Workman JL (2007) Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. *Science* 316(5827):1050–1054
96. Ruan C, Cui H, Lee CH, Li S, Li B (2016) Homodimeric PHD-domain containing Rco1 constitutes a critical interaction hub within the Rpd3S histone deacetylase complex. *J Biol Chem* 291:5428–5438
97. McDaniel SL, Fligor JE, Ruan C, Cui H, Bridgers JB, DiFiore JV, Guo AH, Li B, Strahl BD (2016) Combinatorial histone readout by the dual plant homeodomain (PHD) fingers of Rco1 mediates Rpd3S chromatin recruitment and the maintenance of transcriptional fidelity. *J Biol Chem* 291(28):14796–14802
98. Drouin S, Laramée L, Jacques PE, Forest A, Bergeron M, Robert F (2010) DSIF and RNA polymerase II CTD phosphorylation coordinate the recruitment of Rpd3S to actively transcribed genes. *PLoS Genet* 6(10):e1001173
99. Govind CK, Qiu H, Ginsburg DS, Ruan C, Hofmeyer K, Hu C, Swaminathan V, Workman JL, Li B, Hinnebusch AG (2010) Phosphorylated Pol II CTD recruits multiple HDACs, including Rpd3C(S), for methylation-dependent deacetylation of ORF nucleosomes. *Mol Cell* 39(2):234–246
100. Ruan C, Lee CH, Cui H, Li S, Li B (2015) Nucleosome contact triggers conformational changes of Rpd3S driving high-affinity H3K36me nucleosome engagement. *Cell Rep* 10(2):204–215
101. Huh JW, Wu J, Lee CH, Yun M, Gilada D, Brautigam CA, Li B (2012) Multivalent di-nucleosome recognition enables the Rpd3S histone deacetylase complex to tolerate decreased H3K36 methylation levels. *EMBO J* 31(17):3564–3574

102. Rayasam GV, Wendling O, Angrand PO, Mark M, Niederreither K, Song L, Lerouge T, Hager GL, Chambon P, Losson R (2003) NSD1 is essential for early post-implantation development and has a catalytically active SET domain. *EMBO J* 22(12):3153–3163
103. Angrand PO, Apiou F, Stewart AF, Dutrillaux B, Losson R, Chambon P (2001) NSD3, a new SET domain-containing gene, maps to 8p12 and is amplified in human breast cancer cell lines. *Genomics* 74(1):79–88
104. Eram MS, Kuznetsova E, Li F, Lima-Fernandes E, Kennedy S, Chau I, Arrowsmith CH, Schapira M, Vedadi M (2015) Kinetic characterization of human histone H3 lysine 36 methyltransferases, ASH1L and SETD2. *Biochim Biophys Acta* 1850(9):1842–1848
105. Sun XJ, Wei J, Wu XY, Hu M, Wang L, Wang HH, Zhang QH, Chen SJ, Huang QH, Chen Z (2005) Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. *J Biol Chem* 280(42):35261–35271
106. Li J, Duns G, Westers H, Sijmons R, van den Berg A, Kok K (2016) SETD2: an epigenetic modifier with tumor suppressor functionality. *Oncotarget* 7:50719–50734
107. Jelinic P, Pellegrino J, David G (2011) A novel mammalian complex containing Sin3B mitigates histone acetylation and RNA polymerase II progression within transcribed loci. *Mol Cell Biol* 31(1):54–62
108. Kumar GS, Chang W, Xie T, Patel A, Zhang Y, Wang GG, David G, Radhakrishnan I (2012) Sequence requirements for combinatorial recognition of histone H3 by the MRG15 and Pf1 subunits of the Rpd3S/Sin3S corepressor complex. *J Mol Biol* 422(4):519–531
109. Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T (2010) Regulation of alternative splicing by histone modifications. *Science* 327(5968):996–1000
110. Wen H, Li Y, Xi Y, Jiang S, Stratton S, Peng D, Tanaka K, Ren Y, Xia Z, Wu J et al (2014) ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. *Nature* 508:263–268
111. Guo R, Zheng L, Park JW, Lv R, Chen H, Jiao F, Xu W, Mu S, Wen H, Qiu J et al (2014) BS69/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing. *Mol Cell* 56(2):298–310
112. Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, Akalin A, Schubeler D (2015) Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* 520(7546):243–247
113. Morselli M, Pastor WA, Montanini B, Nee K, Ferrari R, Fu K, Bonora G, Rubbi L, Clark AT, Ottonello S et al (2015) In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse. *eLife* 4:e06205
114. Cai L, Rothbart SB, Lu R, Xu B, Chen WY, Tripathy A, Rockowitz S, Zheng D, Patel DJ, Allis CD et al (2013) An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. *Mol Cell* 49(3):571–582
115. Ballare C, Lange M, Lapinaite A, Martin GM, Morey L, Pascual G, Liefke R, Simon B, Shi Y, Gozani O et al (2012) Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. *Nat Struct Mol Biol* 19(12):1257–1265
116. Brien GL, Gambero G, O'Connell DJ, Jerman E, Turner SA, Egan CM, Dunne EJ, Jurgens MC, Wynne K, Piao L et al (2012) Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation. *Nat Struct Mol Biol* 19(12):1273–1281
117. Fang R, Chen F, Dong Z, Hu D, Barbera AJ, Clark EA, Fang J, Yang Y, Mei P, Rutenberg M et al (2013) LSD2/KDM1B and its cofactor NPAC/GLYR1 endow a structural and molecular model for regulation of H3K4 demethylation. *Mol Cell* 49(3):558–570
118. Pradeepa MM, Sutherland HG, Ule J, Grimes GR, Bickmore WA (2012) Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLoS Genet* 8(5):e1002717
119. Li F, Mao G, Tong D, Huang J, Gu L, Yang W, Li GM (2013) The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSalpha. *Cell* 153(3):590–600
120. Kaplan CD, Laprade L, Winston F (2003) Transcription elongation factors repress transcription initiation from cryptic sites. *Science* 301(5636):1096–1099
121. Neil H, Malabat C, d'Aubenton-Carafa Y, Xu Z, Steinmetz LM, Jacquier A (2009) Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* 457(7232):1038–1042
122. Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Munster S, Cambong J, Guffanti E, Stutz F, Huber W, Steinmetz LM (2009) Bidirectional promoters generate pervasive transcription in yeast. *Nature* 457(7232):1033–1037
123. Li B, Gogol M, Carey M, Pattenden SG, Seidel C, Workman JL (2007) Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. *Genes Dev* 21(11):1422–1430
124. Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Regnault B, Devaux F, Namane A, Seraphin B et al (2005) Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* 121(5):725–737
125. van Dijk EL, Chen CL, d'Aubenton-Carafa Y, Gourvenec S, Kwapisz M, Roche V, Bertrand C, Silvain M, Legoix-Ne P, Loeillet S et al (2011) XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature* 475(7354):114–117
126. Venkatesh S, Li H, Gogol MM, Workman JL (2016) Selective suppression of antisense transcription by Set2-mediated H3K36 methylation. *Nat Commun* 7:13610
127. Churchman LS, Weissman JS (2011) Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* 469(7330):368–373
128. Neri F, Rapelli S, Krepelova A, Incarnato D, Parlato C, Basile G, Maldotti M, Anselmi F, Oliviero S (2017) Intragenic DNA methylation prevents spurious transcription initiation. *Nature* 543(7643):72–77
129. Silva AC, Xu X, Kim HS, Fillingham J, Kislinger T, Mennella TA, Keogh MC (2012) The replication-independent histone H3-H4 chaperones HIR, ASF1, and RTT106 co-operate to maintain promoter fidelity. *J Biol Chem* 287(3):1709–1718
130. Lenstra TL, Benschop JJ, Kim T, Schulze JM, Brabers NA, Margaritis T, van de Pasch LA, van Heesch SA, Brok MO, Groot Koerkamp MJ et al (2011) The specificity and topology of chromatin interaction pathways in yeast. *Mol Cell* 42(4):536–549
131. Huber F, Bunina D, Gupta I, Khmelinskii A, Meurer M, Theer P, Steinmetz LM, Knop M (2016) Protein abundance control by non-coding antisense transcription. *Cell Rep* 15(12):2625–2636
132. Kim JH, Lee BB, Oh YM, Zhu C, Steinmetz LM, Lee Y, Kim WK, Lee SB, Buratowski S, Kim T (2016) Modulation of mRNA and lncRNA expression dynamics by the Set2-Rpd3S pathway. *Nat Commun* 7:13534
133. Sen P, Dang W, Donahue G, Dai J, Dorsey J, Cao X, Liu W, Cao K, Perry R, Lee JY et al (2015) H3K36 methylation promotes longevity by enhancing transcriptional fidelity. *Genes Dev* 29(13):1362–1376



134. Hu Z, Chen K, Xia Z, Chavez M, Pal S, Seol JH, Chen CC, Li W, Tyler JK (2014) Nucleosome loss leads to global transcriptional up-regulation and genomic instability during yeast aging. *Genes Dev* 28(4):396–408
135. Hamilton B, Dong Y, Shindo M, Liu W, Odell I, Ruvkun G, Lee SS (2005) A systematic RNAi screen for longevity genes in *C. elegans*. *Genes Dev* 19(13):1544–1555
136. Pu M, Ni Z, Wang M, Wang X, Wood JG, Helfand SL, Yu H, Lee SS (2015) Trimethylation of Lys36 on H3 restricts gene expression change during aging and impacts life span. *Genes Dev* 29(7):718–731
137. Simon JM, Hacker KE, Singh D, Brannon AR, Parker JS, Weiser M, Ho TH, Kuan PF, Jonasch E, Furey TS et al (2014) Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects. *Genome Res* 24(2):241–250
138. Sorenson MR, Jha DK, Ucles SA, Flood DM, Strahl BD, Stevens SW, Kress TL (2016) Histone H3K36 methylation regulates pre-mRNA splicing in *Saccharomyces cerevisiae*. *RNA Biol* 13(4):412–426
139. Jha DK, Strahl BD (2014) An RNA polymerase II-coupled function for histone H3K36 methylation in checkpoint activation and DSB repair. *Nat Commun* 5:3965
140. Pai CC, Deegan RS, Subramanian L, Gal C, Sarkar S, Blakley EJ, Walker C, Hulme L, Bernhard E, Codlin S et al (2014) A histone H3K36 chromatin switch coordinates DNA double-strand break repair pathway choice. *Nat Commun* 5:4091
141. Winsor TS, Bartkowiak B, Bennett CB, Greenleaf AL (2013) A DNA damage response system associated with the phosphoCTD of elongating RNA polymerase II. *PLoS ONE* 8(4):e60909
142. Carvalho S, Vitor AC, Sridhara SC, Martins FB, Raposo AC, Desterro JM, Ferreira J, de Almeida SF (2014) SETD2 is required for DNA double-strand break repair and activation of the p53-mediated checkpoint. *eLife* 3:e02482
143. Kanu N, Gronroos E, Martinez P, Burrell RA, Yi Goh X, Bartkova J, Maya-Mendoza A, Mistrik M, Rowan AJ, Patel H et al (2015) SETD2 loss-of-function promotes renal cancer branched evolution through replication stress and impaired DNA repair. *Oncogene* 34(46):5699–5708
144. Pfister SX, Ahrabi S, Zalmas LP, Sarkar S, Aymard F, Bachrati CZ, Helleday T, Legube G, La Thangue NB, Porter AC et al (2014) SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Rep* 7(6):2006–2018
145. Qin S, Guo Y, Xu C, Bian C, Fu M, Gong S, Min J (2013) Tudor domains of the PRC2 components PHF1 and PHF19 selectively bind to histone H3K36me3. *Biochem Biophys Res Commun* 430(2):547–553
146. Dhayalan A, Rajavelu A, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, Jeltsch A (2010) The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J Biol Chem* 285(34):26114–26120
147. Rondelet G, Dal Maso T, Willems L, Wouters J (2016) Structural basis for recognition of histone H3K36me3 nucleosome by human de novo DNA methyltransferases 3A and 3B. *J Struct Biol* 194(3):357–367
148. Jin B, Ernst J, Tiedemann RL, Xu H, Sureshchandra S, Kellis M, Dalton S, Liu C, Choi JH, Robertson KD (2012) Linking DNA methyltransferases to epigenetic marks and nucleosome structure genome-wide in human tumor cells. *Cell Rep* 2(5):1411–1424
149. Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 13(7):484–492
150. Park IY, Powell RT, Tripathi DN, Dere R, Ho TH, Blasius TL, Chiang YC, Davis IJ, Fahey CC, Hacker KE et al (2016) Dual chromatin and cytoskeletal remodeling by SETD2. *Cell* 166(4):950–962
151. Zhang X, Huang Y, Shi X (2015) Emerging roles of lysine methylation on non-histone proteins. *Cell Mol Life Sci* 72(22):4257–4272
152. Stabell M, Larsson J, Aalen RB, Lambertsson A (2007) Drosophila dSet2 functions in H3-K36 methylation and is required for development. *Biochem Biophys Res Commun* 359(3):784–789
153. McKay DJ, Klusza S, Penke TJ, Meers MP, Curry KP, McDaniel SL, Malek PY, Cooper SW, Tatomer DC, Lieb JD et al (2015) Interrogating the function of metazoan histones using engineered gene clusters. *Dev Cell* 32(3):373–386
154. Hu M, Sun XJ, Zhang YL, Kuang Y, Hu CQ, Wu WL, Shen SH, Du TT, Li H, He F et al (2010) Histone H3 lysine 36 methyltransferase Hypb/Setd2 is required for embryonic vascular remodeling. *Proc Natl Acad Sci USA* 107(7):2956–2961
155. Zhang Y, Xie S, Zhou Y, Xie Y, Liu P, Sun M, Xiao H, Jin Y, Sun X, Chen Z et al (2014) H3K36 histone methyltransferase Setd2 is required for murine embryonic stem cell differentiation toward endoderm. *Cell Rep* 8(6):1989–2002
156. Dalglish GL, Furge K, Greenman C, Chen L, Bignell G, Butler A, Davies H, Edkins S, Hardy C, Latimer C et al (2010) Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* 463(7279):360–363
157. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, Martinez P, Matthews N, Stewart A, Tarpey P et al (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 366(10):883–892
158. Varela I, Tarpey P, Raine K, Huang D, Ong CK, Stephens P, Davies H, Jones D, Lin ML, Teague J et al (2011) Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature* 469(7331):539–542
159. Mar BG, Bullinger LB, McLean KM, Grauman PV, Harris MH, Stevenson K, Neuberg DS, Sinha AU, Sallan SE, Silverman LB et al (2014) Mutations in epigenetic regulators including SETD2 are gained during relapse in paediatric acute lymphoblastic leukaemia. *Nat Commun* 5:3469
160. Zhu X, He F, Zeng H, Ling S, Chen A, Wang Y, Yan X, Wei W, Pang Y, Cheng H et al (2014) Identification of functional cooperative mutations of SETD2 in human acute leukemia. *Nat Genet* 46(3):287–293
161. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES, Getz G (2014) Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 505(7484):495–501
162. Grosso AR, Leite AP, Carvalho S, Matos MR, Martins FB, Vitor AC, Desterro JM, Carmo-Fonseca M, de Almeida SF (2015) Pervasive transcription read-through promotes aberrant expression of oncogenes and RNA chimeras in renal carcinoma. *eLife* 4:e09214
163. Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K, Sturm D, Fontebasso AM, Quang DA, Tonjes M et al (2012) Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* 482(7384):226–231
164. Wu G, Broniscer A, McEachron TA, Lu C, Paugh BS, Becksfors J, Qu C, Ding L, Huether R, Parker M et al (2012) Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat Genet* 44(3):251–253
165. Behjati S, Tarpey PS, Presneau N, Scheipl S, Pillay N, Van Loo P, Wedge DC, Cooke SL, Gundem G, Davies H et al (2013) Distinct H3F3A and H3F3B driver mutations define

- chondroblastoma and giant cell tumor of bone. *Nat Genet* 45(12):1479–1482
166. Shah MA, Denton EL, Arrowsmith CH, Lupien M, Schapira M (2014) A global assessment of cancer genomic alterations in epigenetic mechanisms. *Epigenetics Chromatin* 7(1):29
167. Fang D, Gan H, Lee JH, Han J, Wang Z, Riester SM, Jin L, Chen J, Zhou H, Wang J et al (2016) The histone H3.3K36M mutation reprograms the epigenome of chondroblastomas. *Science* 352(6291):1344–1348
168. Lu C, Jain SU, Hoelper D, Bechet D, Molden RC, Ran L, Murphy D, Venneti S, Hameed M, Pawel BR et al (2016) Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science* 352(6287):844–849
169. Venneti S, Garimella MT, Sullivan LM, Martinez D, Huse JT, Heguy A, Santi M, Thompson CB, Judkins AR (2013) Evaluation of histone 3 lysine 27 trimethylation (H3K27me3) and enhancer of Zest 2 (EZH2) in pediatric glial and glioneuronal tumors shows decreased H3K27me3 in H3F3A K27M mutant glioblastomas. *Brain Pathol* 23(5):558–564
170. Lewis PW, Muller MM, Koletsky MS, Cordero F, Lin S, Banaszynski LA, Garcia BA, Muir TW, Becher OJ, Allis CD (2013) Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science* 340(6134):857–861
171. Chan KM, Fang D, Gan H, Hashizume R, Yu C, Schroeder M, Gupta N, Mueller S, James CD, Jenkins R et al (2013) The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. *Genes Dev* 27(9):985–990
172. Justin N, Zhang Y, Tarricone C, Martin SR, Chen S, Underwood E, De Marco V, Haire LF, Walker PA, Reinberg D et al (2016) Structural basis of oncogenic histone H3K27M inhibition of human polycomb repressive complex 2. *Nat Commun* 7:11316
173. Yang S, Zheng X, Lu C, Li GM, Allis CD, Li H (2016) Molecular basis for oncohistone H3 recognition by SETD2 methyltransferase. *Genes Dev* 30(14):1611–1616
174. Bjerke L, Mackay A, Nandhabalan M, Burford A, Jury A, Popov S, Bax DA, Carvalho D, Taylor KR, Vinci M et al (2013) Histone H3.3 mutations drive pediatric glioblastoma through upregulation of MYCN. *Cancer Discov* 3(5):512–519