REVIEW ARTICLE | FOCUS
https://doi.org/10.1038/s41594-019-0298-7
nature structural & molecular biology

Histone lysine methyltransferases in biology and disease

Dylan Husmann and Or Gozani*

The precise temporal and spatial coordination of histone lysine methylation dynamics across the epigenome regulates virtually all DNA-templated processes. A large number of histone lysine methyltransferase (KMT) enzymes catalyze the various lysine methylation events decorating the core histone proteins. Mutations, genetic translocations and altered gene expression involving these KMTs are frequently observed in cancer, developmental disorders and other pathologies. Therapeutic compounds targeting specific KMTs are currently being tested in the clinic, although overall drug discovery in the field is relatively underdeveloped. Here we review the biochemical and biological activities of histone KMTs and their connections to human diseases, focusing on cancer. We also discuss the scientific and clinical challenges and opportunities in studying KMTs.

An important mechanism for regulating chromatin involves the reversible covalent modification of histones by chemical moieties such as methyl and acetyl groups. These different chemical 'marks' on histones are linked to discrete chromatin states and regulate the accessibility of DNA to *trans*-acting factors that mediate a wide variety of chromatin-templated processes such as transcription, DNA repair and DNA replication[1]. Chemically, lysine methylation entails the addition of one, two or three methyl groups to the ε-nitrogen of a lysine side chain, forming mono-, di- and trimethylated derivatives (referred to here as Kme1, Kme2, and Kme3, respectively; Fig. 1a). This reaction, while only subtly changing the primary structure of the modified polypeptide, greatly increases the information encoded within the molecule, a feature highlighted by the unique activities frequently coupled to the specific extent of methylation. Methylation of lysines on histone and non-histone proteins is generated by protein lysine methyltransferases (KMTs; referred to as 'writers') and removed by protein lysine demethylases (KDMs; referred to as 'erasers') (Fig. 1a). In the human genome, there are predicted to be over 100 KMTs, and mass spectrometry–based studies suggest that more than 1,000 proteins in the human proteome harbor lysine methylation[2–4].

Lysine methylation was first described in 1959 on a bacterial flagellar protein[5] and soon thereafter identified on histone proteins[6]. Indeed, the core histones contain numerous evolutionarily conserved lysine residues that are methylated in vivo. In humans, the canonical lysine methylation sites are found on histone H3 at lysine 4 (H3K4), lysine 9 (H3K9), lysine 27 (H3K27), lysine 36 (H3K36) and lysine 79 (H3K79), and on histone H4 at lysine 20 (H4K20). These modifications regulate an array of chromatin functions (Fig. 1b)[1]. In addition to these canonical sites, there are several less well characterized sites of lysine methylation on the core histones (for example, H3K23me, H3K63me3, H45me1 and H4K12me1) (Fig. 1c)[4,7]. Together, the substantial numbers of methylation sites and differentially methylated states present in histones illustrate the potential complexity that this signaling system can provide in the regulation of chromatin biology and how its deregulation can lead to disease.

In 2000, the discovery of SUV39H1, the first known histone KMT[8], was a major breakthrough in the field that revealed a direct connection between histone methylation and a classic chromatin-mediated epigenetic phenomenon in flies known as position-effect variegation (PEV)[8] (for a detailed review of PEV, see ref. [9]). Over the past two decades, the discovery and characterization of many additional histone KMTs has uncovered an elaborate network connecting chromatin regulation, epigenetic processes and human disease. In this context, the majority of research on lysine methylation has naturally focused on histone substrates and its role in chromatin and epigenetic regulation. One unintended consequence of this emphasis has been the emergence of biases in the initial characterization of the catalytic activities of orphan KMTs as histone-modifying enzymes. For example, the availability of reagents such as state-specific antibodies with which to study histone methylation, combined with the potential underappreciation of the limitations of these reagents, has led to the mischaracterization of some enzymes as histone KMTs[2] (discussed below). As the correct assignment of catalytic specificity for KMTs is crucial for understanding the role of chromatin in disease and for efforts to develop therapeutics, here we offer our perspective in classifying the reported histone KMTs as (1) bona fide histone-modifying enzymes, (2) enzymes that are referenced in the literature as histone KMTs but clearly are not, or (3) enzymes for which further work is necessary before any meaningful conclusions about catalytic activity and specificity can be drawn. Our rationale for making these distinctions, and their implications for disease etiology, are discussed below.

KMTs that catalyze canonical histone lysine methylation

In the human proteome, there are two domains with annotated lysine methyltransferase activity: the SET domain (named for three *Drosophila melanogaster* proteins originally recognized as containing the domain: Su(var)3–9, enhancer of zeste and trithorax) and the seven-beta-strand (7βS) domain (which is found on enzymes ranging from the histone KMT hDOT1L (Fig. 2a) to DNA methyltransferases)[1–3]. In humans, there are 55 SET-domain-containing proteins. Of these, half are active KMTs (methylating histone and/or non-histone substrates), one protein (SETD3) is a histidine methyltransferase[10], and the enzymatic activities of the remainder are unclear[2] (Fig. 2a,b and Table 1). The 7βS family is larger and more diverse than the SET family, with approximately 150–160 members in humans[3,11]. Different 7βS-containing proteins methylate a wide range of substrates including lysine, arginine, other amino acid side chains, N-terminal α-amines, DNA, RNA and various metabolites[3].

The canonical histone lysine methylation marks found in humans are shown in Fig. 1b. These various modifications are generated in a context-dependent manner by a total of 24 different enzymes: 23 different SET proteins and one 7βS protein (Fig. 2a). In general, histone KMTs are highly selective: that is, the enzymes that methylate H3K36 do not methylate a different lysine if K36 is mutated. One exception is the meiotic recombination factor PRDM9, which

Department of Biology, Stanford University, Stanford, CA, USA. *e-mail: ogozani@stanford.edu

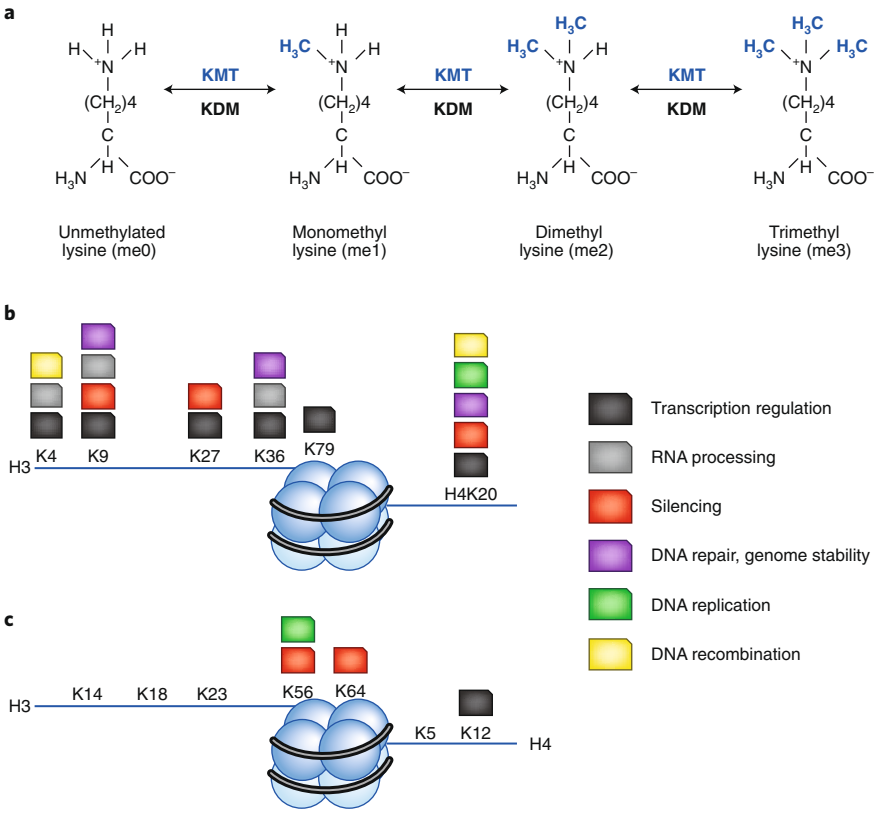

Fig. 1 | Main sites of lysine methylation on mammalian histones and chromatin functions. a, Chemical structures of methylated derivatives of lysine. Lysine residues can be monomethylated, dimethylated or trimethylated. **b**,**c**, Canonical (**b**) and non-canonical (**c**) lysine methylation marks on core nucleosomal histone H3 and H4 and their basic functions. Numbers adjacent to 'K' indicate the positions of the methylated lysines on histone H3 or histone H4. DNA is shown as black lines wrapped around blue histones. Key (right), chromatin-related functions associated with the methylation at left.

trimethylates H3K4 in vivo in meiotic cells, but in vitro also methylates H3K9 and H3K36 (Fig. 2a,b). The physiological importance of the H3K9 and H3K36 activities of PRDM9 remains to be determined. The 7βS protein hDOT1L is the only enzyme in the human proteome that generates H3K79me, one of the few histone modifications found within the globular region of the nucleosome (Fig. 1b). In contrast, multiple enzymes mediate methylation events at H3K4, H3K9 and H3K36 (Fig. 2a). This enzymatic redundancy is used for targeting specific activities in a context-dependent manner such as differential genomic localization (such as methylation at an enhancer versus promoter region) and for the selective generation of different methylation states (such as me2 versus me3).

For example, H3K36me2 is generated by four related enzymes (NSD1, NSD2, NSD3 and ASH1L), whereas SETD2 is the only enzyme in somatic cells that synthesizes H3K36me3 (Fig. 2C). Notably, the generation of H3K36me3 by SETD2 is not dependent on the presence of H3K36me2[12,13]; i.e., the initial recognition of the nucleosome as a substrate by SETD2 is far more efficient on unmethylated H3K36 than H3K36me2 (Fig. 2c). At H4K20, the monomethylated state is generated solely by SETD8, and the higher methylation states are synthesized by the KMTs SUV420-H1 and SUV420-H2 (Fig. 2c)[14]. However, unlike SETD2, SUV420-H1 and SUV420-H2 prefer a methylated substrate (H4K20me1) to unmethylated H4K20[14,15] (Fig. 2c). As a consequence, deletion of *SETD8* leads to loss of all H4K20 methylation states even though SETD8 generates only the monomethyl species[14–16].

Notably, SETD8 and several other KMTs that methylate histones also modify non-histone substrates[2]. For instance, SETD7, G9A, GLP and SETD8 methylate p53 (as well as other non-histone substrates)[2]. In this context, knockout of *Setd8* in *Drosophila* is lethal, whereas flies harboring a substitution of H4K20A, which prevents methylation of this residue, have a substantial delay in development but are otherwise normal[17]. The more severe phenotype that results from the *Setd8* deletion versus the H4K20A mutation argues for physiologically important roles of SETD8 outside of H4K20 methylation. Thus, for select histone KMTs, their ability to methylate non-histone substrates must be taken into account in evaluating potential inhibitory compounds as candidate therapeutics.

Considerable efforts have been made to develop small-molecule inhibitors of different histone KMTs as tool compounds and for therapeutic purposes[18]. At present, active clinical trials (phase 1 and 2) are focused on several inhibitors of EZH2 (the main H3K27 KMT) and one inhibitor of the essential EZH2 cofactor EED; these compounds are being evaluated for efficacy in the treatment of a wide range of adult and pediatric neoplasm types (for example, ClinicalTrials.gov identifiers NCT03456726, NCT03213665, NCT02900651 and NCT02395601). Patients enrolled in the EZH2/EED inhibitor trials have tumors that share a common molecular signature: they either are positive for *EZH2* gain-of-function mutations or harbor loss-of-function mutations in other chromatin-regulatory factors that are predicted to create cellular dependency on EZH2 activity. Beyond EZH2, a clinical compound targeting hDOT1L was evaluated in a phase 1 trial that was completed in 2016 (ClinicalTrials.gov identifier NCT02141828), but as of this writing a phase 2 trial has not commenced. Tool and preclinical compounds also exist for several other histone KMTs (for example, SETD8 and G9A)[18], arguing that KMTs, as an enzyme class, are druggable. However, several obstacles need to be overcome in developing drugs against some of the more promising KMT targets, including the lack of structural information about the enzymes, the need to

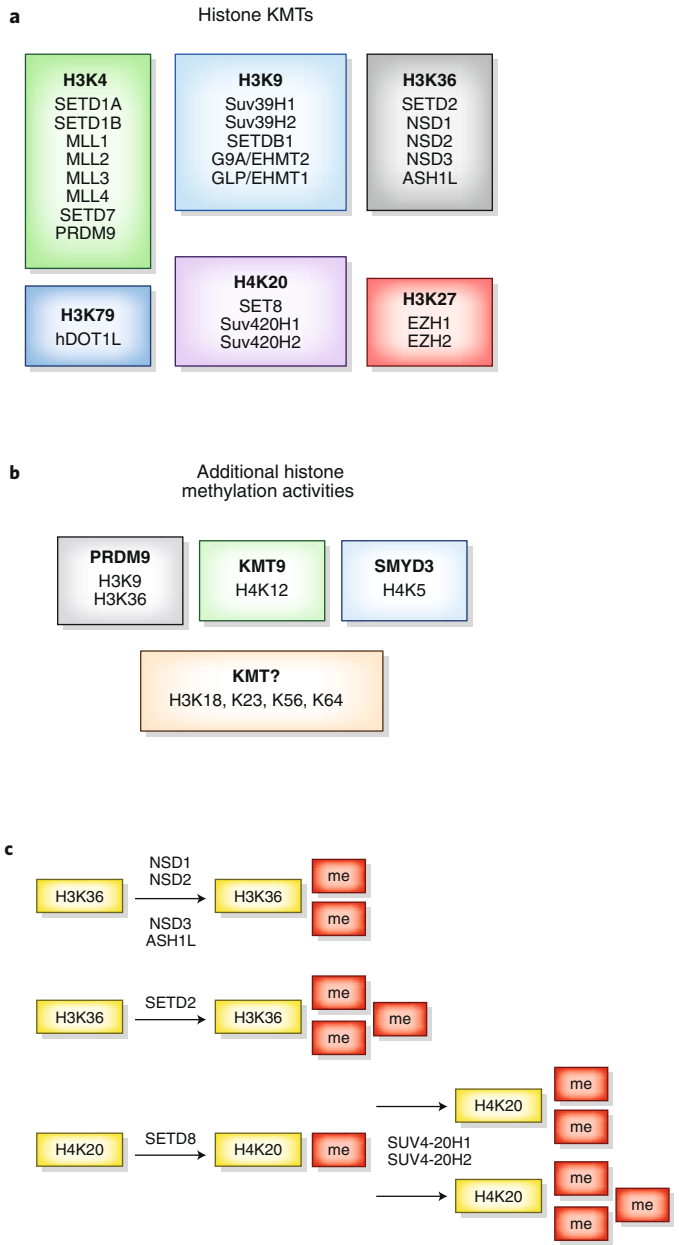

Fig. 2 | Histone KMTs in the human proteome. a, Human histone KMTs categorized by their established substrate specificity. **b**, Top: Examples of additional histone KMT activities. Bottom: Methylation is also detected at the non-canonical H3K18, K23, K56 and K64 sites, but the enzymes catalyzing these events are not known. **c**, Top two rows: generation of H3K36 trimethylation is not dependent on existing dimethylation. Bottom row: generation of H4K20me2 and H4K20me3 is dependent upon SETD8-generated H4K20me1.

use nucleosomes as substrates for in vitro drug screening, and the still limited, although growing, understanding of the types of compounds best suited to engage KMTs.

Mistaken identity: not all KMTs methylate histones

In addition to the enzymes listed in Fig. 2a,b, the following candidate KMTs—many of which are linked to human disease—have been reported to generate at least one canonical histone methylation mark (the putative modified residue(s) is (are) provided in parentheses after each enzyme symbol): MLL5 (H3K4), SETD3 (H3K4, H3K36), SETD5 (H3K9), SETDB2 (H3K9), SMYD1 (H3K4), SMYD2 (H3K36), SMYD3 (H3K4), SMYD5 (H4K20), SETMAR (H3K36), PRDM1 (H3K9), PRDM2 (H3K9), PRDM3 (H3K9), PRDM8 (H3K9) and PRDM16 (H3K4, H3K9) (Fig. 2a). Of these enzymes, the specific canonical histone methylation activities reported for MLL5, SETD3, SETDB2, SMYD3 and SETMAR have been independently tested and not reproduced[10,19–22]; the original report on MLL5 was retracted[23]. Moreover, biophysical and biochemical analyses of MLL5 indicate that it is not an active enzyme[22]. SETD5, an important protein etiologically linked to intellectual disability disorders[24,25], is similar in structure to MLL5 and therefore is not likely to be an active enzyme[22]. We recently demonstrated that SETD3 is a highly selective histidine methyltransferase and that it has no detectable activity on nucleosomes[10] (Fig. 2d). SETDB2, given its sequence similarity to SETDB1, is assumed to be an H3K9 methyltransferase; however, to date, no activity for SETDB2 has been rigorously identified[21]. SMYD3 is a largely cytoplasmic protein that methylates non-histone substrates such as MAP3K2 and does not methylate H3K4 on free histones or on nucleosomes[20]. SETMAR is a DNA-repair protein that consists of a fusion between a SET domain and a DNA transposase domain[19,26]. In vitro, SETMAR methylates free H3 and H2B but has no activity on nucleosomes, and its activity on free H3 does not target K36, as determined by tandem mass spectrometry[19]. Thus, the physiological substrate of SETMAR and its potential role in DNA repair remains to be elucidated. SMYD2 is a relatively promiscuous enzyme (as far as KMTs go) and methylates many substrates. However, it has no activity on nucleosomes and lacks specificity on free histones, in contrast to its interaction with p53, one of SMYD2's better-characterized substrates, where it shows a highly selective activity[27]. Zebrafish SMYD1 has activity toward histones[28], but methylation of histones has not been demonstrated for the human homolog. There is one report showing that SMYD5 has H4K20 trimethylation activity[29]. However, deletion of both *Suv420-H1/2* in mice eliminates H4K20me3[14], leaving the status of SMYD5 as a bona fide H4K20-modifying enzyme unresolved. Finally, several conflicting reports have suggested that PRDM1, PRDM2, PRDM3, PRDM8 and PRDM16 methylate H3K9 or H3K4, but other researchers have been unable to reproduce such activities, and thus more definitive work is required to determine whether

Table 1 | Putative KMTs reported in the literature to be specific histone-modifying enzymes compared to their actual activity on histones and other substrates

Putative KMT	Histone physiological substrate	Physiological substrate(s)
SETDB2[21]	?	?
SETMAR[19,26]	?/–*	?
SETD3[10,30]	–*	Actin-H73
SETD5[22,24]	?	?
MLL5[22]	–	–
SMYD1[28]	?	?
SMYD2[27,94]	–*	p53, MAPKAP3,
SMYD3[20,95,96]	H4K5me	MAP3K2, VEGFR1
SMYD5[29]	H4K20me3?	?
PRDM1[31]	H3K9me?	?
PRDM2[31]	H3K9me?	?
PRDM3[97]	H3K9me?	?
PRDM8[31]	–	?
PRDM16[97,98]	H3K4me?, H3K9me?	?

?, unknown; –, no methylation activity; *, no histone methylation activity on nucleosomes; histone site followed by "?", more evidence required to determine whether the reported activity is reproducible.

these biologically important proteins are truly active enzymes, and if they are, to identify their physiological substrates[2] (Fig. 2d).

Why have these enzymes been potentially mischaracterized? As mentioned above, in some cases, interpretation of data relying solely on state-specific antibodies can be misleading, particularly in the absence of tandem mass spectrometry studies as an independent approach to confirm the specific methylation event. In addition, although technically challenging, the use of nucleosomes as substrates in addition to histone peptides can provide important information about whether a putative histone methylation activity is likely to be physiologically relevant. Finally, some studies have relied on mass spectrometry data in which the mass shifts attributed to methylation reactions are the wrong molecular weight for methylation, raising doubt about the studies' conclusions (for example, ref. [30]). Taken together, these complications emphasize that it is important that those undertaking any research, including drug-development efforts, focused on these potentially mischaracterized KMTs—many of which have clear links to disease (for example, PRDM1 and PRDM2 are important tumor suppressors[31], PRDM16 is a key regulator of adipogenesis[31] and SETDB2 is a regulator of fibrotic diseases[21])—consider starting with a rigorous and unbiased analysis of their enzymatic activities.

H3K36-specific lysine methyltransferases in cancer

The established link between histone lysine methylation dynamics, gene expression regulation and oncogenic programming provides a paradigm for the way that pathological alterations of histone KMTs can promote the development and progression of diverse cancers (Fig. 3). The findings in this field are vast, and many excellent and comprehensive reviews on the topic are available for the interested reader[1,32,33]. Here we focus on the pathological roles of the enzymes that either dimethylate or trimethylate H3K36 (see Fig. 2c) as model histone KMTs and discuss examples of crosstalk between H3K27 and H3K36 methylation in epigenetic-mediated oncogenic programming.

The state of methylation at H3K36 defines distinct biological outcomes, and mutations in the H3K36 KMTs are linked to a variety of developmental disorders and cancer (Fig. 3 and Table 2). SETD2, which synthesizes H3K36me3 in humans, regulates DNA methylation, RNA processing, DNA repair and tumor suppression[13,34–37]. In contrast, it is less clear what specific molecular functions are associated with H3K36me2, although this modification has been linked to DNA methylation, gene activation and cellular transformation[12,33,38,39]. There are four enzymes that generate H3K36me2: NSD1, NSD2, NSD3 and ASH1L (Fig. 2c). The enzyme(s) responsible for generating H3K36me1 is (are) unknown, and a cellular function for H3K36me1 is not clear at present, although it is likely that the mark itself is synthesized through the combined actions of KMTs and KDMs.

The initial evidence for a potential tumor-suppressive role of SETD2, and by proxy H3K36me3, came from sequencing studies of renal-cell carcinoma (RCC). These studies found recurrent biallelic loss of *SETD2*, a classic hallmark of known tumor-suppressor genes[40,41]. Subsequent sequencing studies have identified recurrent *SETD2* mutations across a broad spectrum of human malignancies, including lung adenocarcinoma (LUAC)[42], multiple types of leukemia and other hematological malignancies[43–48], central nervous system tumors[49], bladder cancer[50] and gastrointestinal tumors[51]. The remarkably high frequency of *SETD2* mutations across a wide variety of cancer types is reminiscent of other classical tumor suppressors and suggests a broad, general role in prevention of cancer.

Further sequencing efforts in RCC found that inactivating mutations in *SETD2* occurred in a subclonal fraction of the tumor, arguing against a role for *SETD2* loss in tumor initiation. However, the same studies identified individual tumors with distinct inactivating mutations in *SETD2*[52], demonstrating parallel evolution toward loss of *SETD2*. These results suggest that SETD2 loss plays a role in cancer progression. Indeed, biallelic loss of *SETD2* in patients with RCC is unfortunately associated with significantly lower survival rates[53]. Furthermore, although *SETD2* is lost in only a fraction of all RCCs (11.3%), its loss is far more prevalent in more aggressive subtypes of RCC (63%)[54]. The role of SETD2 in cancer progression extends to high-risk gastrointestinal stromal tumors[51]. Finally, patients with acute lymphocytic leukemia (ALL) who unfortunately relapsed after chemotherapy frequently acquire loss-of-function mutations in *SETD2*[55].

More recent analyses of multiple cancer subtypes have provided additional evidence that loss of *SETD2* drives tumor progression. In comprehensive studies of acute myeloid leukemia (AML), ALL and LUAC, *SETD2* was mutated at a much higher frequency in tumors driven by fusion oncogenes[56]. Specifically, mutations in *SETD2* were detected in 22.6% of patients with MLL-rearranged leukemia but in only 4.6% of patients without the oncogenic fusion. Similarly, in LUAC, *SETD2* mutations were found in 18% of patients with cancers driven by fusion oncogenes, compared with 9% of those lacking oncogenic fusions. Other studies have observed co-occurrence of *SETD2* loss and silencing of *CDKN2A*[42], as well as Ras-activating mutations[55]. This co-occurrence suggests that the requirement for SETD2 in tumor suppression may be enhanced in specific contexts.

In vivo screens for tumor suppressors have identified SETD2 as a top candidate in multiple cancer models, including ALL[57], hepatocellular carcinoma[58] and gastrointestinal cancer[59]. Moreover, recent work using multiplexed in vivo CRISPR-based genome editing to knock out numerous known and putative tumor-suppressor genes in a Kras-driven mouse model of LUAC demonstrated that *Setd2* depletion dramatically increased tumor size[60]. Indeed, loss of *Setd2* resulted in the largest tumors observed in the study, surpassed only by tumors harboring p53 inactivation. These results are consistent with earlier work demonstrating a role for SETD2 in suppressing Kras-driven LUAC in mouse models[61]. Collectively, these studies provide compelling experimental evidence for the tumor-suppressor function of SETD2, in accordance with the numerous *SETD2* mutations identified in human tumors (Fig. 3). However, further work is needed to determine the mechanism(s) of tumor suppression by SETD2, the relationship to H3K36me3 catalysis and the effects of genetic context.

At the molecular level, the ability of SETD2 to regulate several fundamental biological processes is directly linked to selective recognition of H3K36me3 by methyl-lysine reader domains. Baubec et al. demonstrated that crosstalk between DNA methylation and histone methylation is mediated by recognition of H3K36me3 by the DNMT3B PWWP domain[37]. This interaction facilitates the targeting of DNMT3B to the bodies of transcribed genes, which are enriched for H3K36me3. This leads to focal de novo DNA methylation at these genomic regions[37], which may influence the expression of nearby genes. Indeed, *SETD2* loss in mouse oocytes causes defects in DNA methylation, genomic imprinting and development[62]. SETD2-mediated regulation of transcription and pre-mRNA splicing is also mediated by an H3K36me3-selective reader domain[35,36]. The tandem bromo-PWWP domains on the nuclear factor BS69 (also known as ZMYND11) recognizes K36me3, but only in the context of nucleosomes containing the histone variant H3.3[36]. The binding of BS69/ZMYND11 to H3.3K36me3-enriched chromatin recruits BS69/ZMYND11 and its associated proteins, which include RNA splicing and transcription factors[35,36]. Connections between DNA-repair mechanisms and SETD2 are also mediated by reader domains. For example, recognition of H3K36me3 by the PWWP-domain-containing factor MSH6 facilitates the association of the mismatch-recognition complex to facilitate DNA repair[63]. Collectively, the discovery of selective H3K36me3-reader domains has provided crucial insight into the molecular mechanisms of action by which SETD2 regulates biology and how these functions may influence oncogenesis. For example, altered gene expression

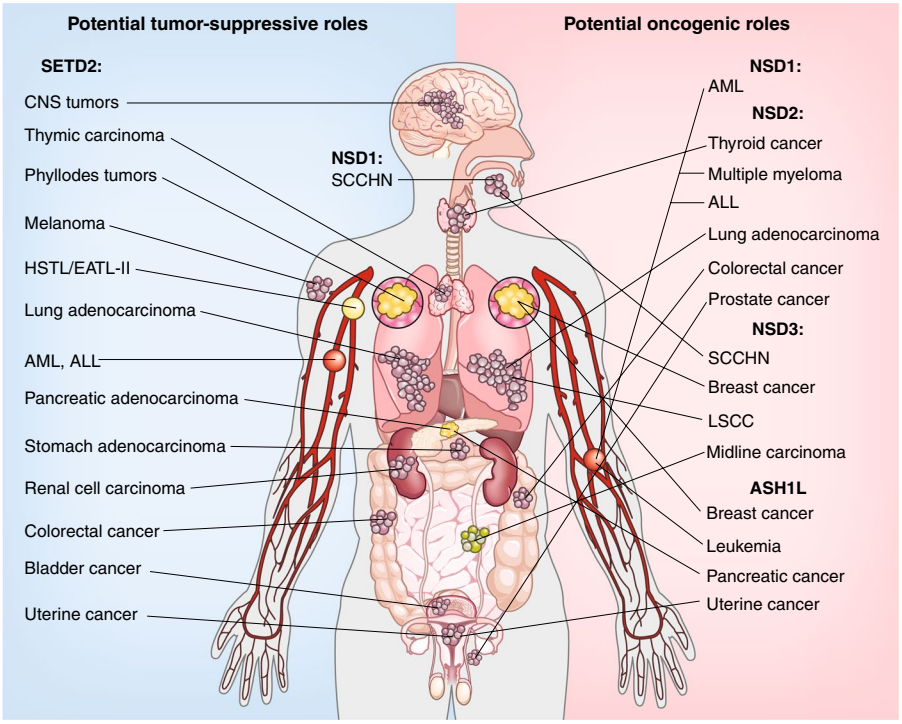

Fig. 3 | Spectrum of cancers associated with H3K36 methyltransferases. The potential tumor-suppressive functions listed are informed by the identification of recurrent deletions, frameshifts, or truncating or damaging missense mutations, and by biological studies, including mouse models. Potential oncogenic functions are informed by overexpression, focal amplifications, gain of function or identification of a fusion oncogene, and by biological studies, including mouse models. CNS, central nervous system; HSTL, hepatosplenic T cell lymphoma; EATL-II, enteropathy-associated T cell lymphoma, type II; LSCC, lung squamous-cell carcinoma.

and/or compromised DNA repair in a SETD2-deficient setting could promote cellular transformation. Finally, two interesting studies recently reported that SETD2 directly methylates tubulin and STAT1[64,65]. The strong substrate preference of SETD2 for intact nucleosomes suggests that it recognizes a specific three-dimensional topology during catalysis. Therefore, the molecular basis for the recognition of such disparate substrates (nucleosomal H3K36, tubulin and STAT1) and the relative contributions of H3K36me3 versus the non-histone substrates in tumor suppression are important questions to address in the future.

Whereas H3K36me3 is generated exclusively by SETD2, the biosynthesis of H3K36me2 is more complex. There are four related enzymes, NSD1, NSD2, NSD3 and ASH1L, that can generate H3K36me2 on nucleosomes in vitro (Fig. 2c)[66]. In most cell types, including various cancer cell lines, NSD2 is responsible for generating the bulk of H3K36me2[12]. In specific cellular contexts NSD1 replaces NSD2 as the enzyme required to generate global H3K36me2[67]. ASH1L does not globally regulate cellular H3K36me2 levels; instead, its activity is localized to specific genes[68]. The physiological role of NSD3 is not clear.

In contrast to SETD2 and its role as a tumor suppressor, all four of the H3K36me2-specific KMTs are thought to promote oncogenesis (Fig. 3). One clear example is the role of NSD2 in the pathogenesis of multiple myeloma (MM)[69]. MM is an incurable blood malignancy that effects hundreds of thousands of people throughout the world[70,71]. Among patients with MM, 15–20% carry a t(4;14) translocation, which places the transcription of *NSD2* under the control of a strong IgH intronic enhancer and leads to aberrant, massive upregulation of NSD2 that is thought to drive cancer development[72–74]. Consistent with this, NSD2 expression in MM cells drives xenograft tumor formation and tumor invasion in mice in a manner that depends on the catalytic activity of NSD2[39]. Beyond MM, NSD2 overexpression is broadly found in diverse cancers[39] and drives metastatic progression in prostate cancer[75]. Consistent with this expression profile, NSD2 depletion in multiple cancer cell lines results in decreased cellular proliferation[12,76].

In addition to the t(4;14) translocation and general overexpression of NSD2, a recurrent heterozygous gain-of-function NSD2 (E1099K) variant is found in ~10% of cases of childhood ALL with a precursor-B phenotype[77,78]. ALL is the most common cancer diagnosed in children, representing more than a quarter of all pediatric neoplasms[79]. E1099 is found within the catalytic SET domain of NSD2, and the E1099K substitution confers a roughly 1.5-fold increase in NSD2 catalytic efficiency through a mechanism that is presently unknown. Expression of NSD2-E1099K in cells leads to elevated H3K36me2 levels, which causes a decrease in H3K27me3 levels due to the direct inhibition of EZH2 by H3K36me2[70,71,80]. In this context, the ability of NSD2-E1099K to drive pediatric ALL is postulated to be mediated in part via depletion of H3K27 methylation, which in turn leads to defects in epigenetic gene silencing and oncogenic reprograming (Fig. 4).

Notably, beyond pediatric ALL, the NSD2 E1099K mutation is found in other neoplasms, including several types of solid tumors, such as LUAC, colon cancer and thyroid tumors[71,73,81]. Together, the many links between NSD2 alterations and different cancers indicate that the NSD2–H3K36me2 axis has a broad role in promoting tumorigenesis. However, it remains unclear whether H3K36me2 has direct effect on chromatin and gene regulation beyond the suppression of H3K27me3. That said, depletion of NSD2 and H3K36me2 in HT1080 cancer cells impairs cell proliferation, and this phenotype is independent of H3K27me3, because it is not rescued by EZH2 inhibition[82]. Moreover, the PWWP domain on NSD2 itself preferentially binds to H3K36me2, and it is postulated to be important for the propagation of NSD2-mediated H3K36me2 domains[83]. This

Table 2 | Histone KMTs from Fig. 2 in human disease, observed murine phenotypes, and putative roles in cancer

	Human disease or syndrome	Knockout mouse phenotype(s)	Putative roles in cancer
H3K4			
SETD7		−/−: no identifiable phenotype[99]	
PRDM9		−/−: infertile; involved in speciation[100]	Overexpressed in specific cancers
SETD1A	Schizophrenia[101]	−/−: lethal	
		Conditional: impaired B-cell development[102]	Oncogenic/tumor suppressive
SETD1B	12q24.31 microdeletion[103]	−/−: lethal, hematopoietic defects[104]	Oncogenic/tumor suppressive
MLL1–MLL4	MLL1: Wiedmann-Steiner[105]	−/−: lethal[106–109] Conditional: pleiotropic, e.g., impaired hematopoiesis in *MLL1*[111]	Mll fusions oncogenic
	MLL2: dystonia[110]		Wild-type Mll
	MLL3: Kleefstra-2[112]		Tumor suppressive
	MLL4: Kabuki[113]		
H3K9			
G9A		−/−: lethal[114]	Oncogenic
		Conditional: impaired learning and memory[115]	
GLP	Kleefstra-1[116]	−/−: lethal[117]	Oncogenic
		Conditional: impaired learning, adipogeneis[115,118]	
SUV39H1, SUV39H2		−/−: no identifiable phenotype	
		dKO: increased tumor risk, impaired fertility[114]	
SETDB1		−/−: lethal; defects in neural development[119]	Oncogenic
		Conditional: impaired spermiogenesis, oogenesis[120]	
H3K27			
EZH1		−/−: no identifiable phenotype[121]	
		Conditional: impaired hematopoeisis[122]	
EZH2	Weaver[89]	−/−: lethal[123]	Oncogenic/tumor suppressive
		Conditional: broadly impaired development	
H3K36			
NSD1	Sotos[88], Beckwith-Wiedmann[125]	Conditional: lethal[124]	Oncogenic/tumor suppressive
NSD2	Wolf-Hirschhorn[126–128]	−/−: lethality shortly after birth	Oncogenic
		-/+: WHS-like defects[129]	
		Conditional: impaired B-Cell development[130]	
NSD3			Oncogenic
ASH1L	Intellectual disability[131]	−/−: lethal[132]	Oncogenic
		Conditional: impaired hematopoesis[133]	
SETD2	Luscan-Lumish[134,135]	−/−: lethal, vascular defects[136]	Tumor suppressive
		Conditional: impaired osteogenesis[137],	
		myogenesis[138], germ cell development[62,139], hematopoesis[140]	
H3K79			
DOT1L		−/−: lethal: impaired cardiac development[141]	Oncogenic
		Conditional: impaired hematopoesis[142,143]	
H4K20			
SETD8		−/−: embryonic lethal[144]	
SUV420H1	Intellectual disability[145]	−/−: lethal shortly after birth; short stature[14]	
SUV420H2		−/−: no apparent defects	
H4K12			
KMT9			Oncogenic

dKO, double knockout; WHS, Wolf-Hirschhorn syndrome.

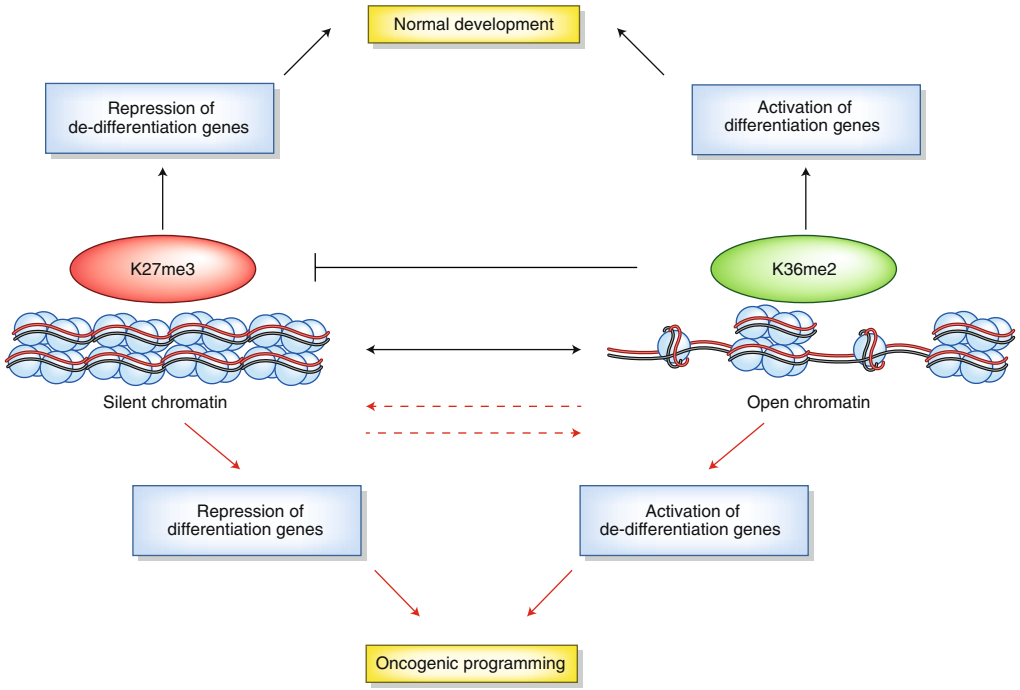

Fig. 4 | Model for crosstalk between methylation at H3K27 and H3K36 in oncogenic programming. Deregulation of the dynamic interplay between methylation at H3K27 and that at H3K36 leads to pathological transcriptional activation or repression and thereby promotes oncogenic reprogramming.

might indicate that other yet-to-be-discovered H3K36me2-specific reader domains exist that link this modification to cancer pathways.

Like NSD2, the proteins NSD1, NSD3 and ASH1L are linked to oncogenesis. In AML, the t(5;11) fusion of NUP98, a member of the nuclear pore complex, to NSD1 (NUP98–NSD1), is found in about 5% of AML cases and is associated with poor prognosis[39]. In mouse adaptive-transfer experiments, Wang et al. showed that bone marrow progenitor cells ectopically expressing NUP98–NSD1 rapidly developed AML[84]. Mechanistically, this transformation activity is mediated by the activation of HOX genes, important developmental genes that are frequently dysregulated in cancer, via H3K36 methylation and antagonism of EHZ2-mediated repression[84]. Beyond the NUP98–NSD1 fusion, the role of NSD1 in cancer is complex. Nonsense mutations in *NSD1* are observed in ~10% of distinct populations of patients with squamous cell carcinoma of the head and neck (SCCHN)[67] and with lower frequency in several other cancers[85]. This suggests that loss of functional NSD1 promotes oncogenesis. On the other hand, patients with SCCHN who harbor *NSD1* mutations have a favorable outcome and show a better response to chemotherapy. Thus, the role of NSD1 in cancer might be dependent on the tissue and etiological context as well as the mutational landscape of the disease.

The *NSD3* gene is commonly amplified in breast cancer, lung squamous cell carcinoma and squamous cell carcinoma of the head and neck[81]. In addition, *NSD3* is involved in rare translocations in patients suffering from acute myeloid leukemia (AML), with the fusion including the NSD3 SET domain[86]. Furthermore, a rare fusion lacking the SET domain but including the BET-interaction domain of NSD3 is found in midline carcinoma[84,87]. Despite the links between *NSD3* and cancer, the physiological role of the catalytic activity of NSD3 and its relationship to tumorigenesis remains unclear.

Like *NSD3*, the gene *ASH1L* is amplified in various cancers including breast, uterine and pancreas[81]. Moreover, the H3K36 dimethylation activity of ASH1L promotes MLL-dependent leukemogenesis in both mouse models and human MLL-rearranged leukemic cells through the regulation of transcription at key leukemia-associated gene[68]. Taken together, these data indicate that whereas the SETD2–H3K36me3 axis plays a role in suppressing tumorigenesis, the dimethyl state at H3K36 is generally associated with promoting oncogenesis.

Although either dimethylation or trimethylation at H3K36 antagonizes EZH2 and prevents H3K27 methylation, the distribution patterns of H3K36me2 and H3K36me3 across the genome are distinct. H3K36me2 is far more abundant than H3K36me3 and is both present in intergenic regions and enriched proximal to the transcriptional start sites of actively transcribed genes[12]. In contrast, H3K36me3 is found largely within the bodies of actively transcribed genes, reaching its highest levels at the 3' end. Overall, it is likely that their different chromatin distributions coupled with state-specific reader domains explain why H3K36me2 and H3K36me3 have divergent roles in cancer. Furthermore, because of the wider distribution of H3K36me2 across the genome and its overall greater abundance, H3K36me2 will naturally have a larger impact in antagonizing EZH2 than does SETD2-catalyzed H3K36me3. It appears contradictory that although EZH2 itself has oncogenic properties, counteracting its activity promotes H3K36me2-driven cancers. This, however, highlights an important concept in epigenetic misregulation: either gains or losses of histone methylation marks in a cell-context-dependent manner can select for gene expression programs that provide a fitness advantage or prevent differentiation, locking cells into a proliferative state that exists only transiently in normal development (Fig. 4).

Histone KMTs and developmental disorders

Haploinsufficiency of histone KMTs manifest in numerous developmental disorders (Table 2). Notably, there is striking similarity in the developmental phenotypes of patients with deficiencies in the main H3K27 methyltransferase complex and those with deficiencies in H3K36 methyltransferases (Table 2). Both Sotos syndrome and Weaver syndrome, largely characterized by mutations in *NSD1* and *EZH2*, respectively, present with overgrowth and intellectual disability[88,89]. Although these conditions are categorized under different names, a subset of patients with Weaver syndrome

possess *NSD1* mutations rather than *EZH2* mutations[90]. Furthermore, patients diagnosed with Sotos syndrome but lacking mutations in *NSD1* have been found to have mutations in *SETD2* and *DNMT3A*[91]. The remarkable phenotypic convergence observed in these patients may reflect underlying molecular relationships among the methylation of H3K27, H3K36 and DNA.

Non-canonical histone methylation sites in disease

Beyond the canonical sites, many other methylation events on histones (for example, H3K14me3, H3K56me1, H3K64me3, H4K12me1 and several others) have been identified by various methods, including mass spectrometry[4,7] (Fig. 1c). Interestingly, H3K14me3 is not normally found in human chromatin but is generated by the bacterial effector protein RomA in cells infected with *Legionella pneumophila*[92]. This mark is hypothesized to repress the expression of host genes encoding components of the innate immune system, which helps promote intracellular replication of the mycobacterium. Thus, a RomA inhibitor could function as an antibiotic to selectively treat legionella pneumonia[92].

Recently, KMT9, a heterodimeric complex consisting of two 7βS enzymes (C21orf127, also known as HEMK2, N6AMT1 or PrmC; and TRMT112), was shown to monomethylate H4K12 in vitro on nucleosomes, and depletion of KMT9 in prostate cancer cells results in decreased endogenous H4K12me1 levels[93]. Beyond DOT1L, KMT9 represents the only other 7βS enzyme known to date to have histone lysine methylation activity. In cells, H4K12me1 modification localizes to gene promoters, and depletion of this mark by knockdown of KMT9 reduces the expression of genes marked with H4K12me1, suggestive of a role for the KMT9–H4K12me1 axis in transcription initiation[93]. Notably, the levels of KMT9 and H4K12me1 are specifically elevated in malignant prostate cancer. Furthermore, depletion of KMT9 impairs cell proliferation and xenograft tumor growth of androgen-independent prostate cancer, but not the growth of several other cell types. Interestingly, the heterodimeric KMT9 complex also functions as a protein glutamine methyltransferase, but Metzger et al. have identified KMT9 mutants that separate the two enzymatic functions to demonstrate H4K12me1 synthesis as the relevant activity in prostate cancer[93]. This study suggests that targeting of a non-canonical histone mark, H4K12me1—through inhibition of KMT9—may offer a new strategy for the treatment of lethal castration-resistant prostate cancer.

Outlook

Over the last several decades, fueled by discoveries based on the integration of diverse methods, the scientific community has developed an understanding of the fundamental role of histone lysine methylation in the regulation of chromatin biology and of how this complex signaling system affects human disease. Drugs targeting EZH2, the main H3K27 KMT, are being tested as precision medicines that will hopefully soon be available in clinical settings to help patients. Over the next several years, we anticipate that new ways to chemically or biologically modulate other histone KMTs, such as NSD2, will be realized and may offer therapeutic benefit in the treatment of cancers and other human pathologies.

Received: 18 July 2019; Accepted: 12 August 2019;
Published online: 3 October 2019

References

1. Murn, J. & Shi, Y. The winding path of protein methylation research: milestones and new frontiers. *Nat. Rev. Mol. Cell Biol.* **18**, 517–527 (2017).
2. Carlson, S. M. & Gozani, O. Nonhistone lysine methylation in the regulation of cancer pathways. *Cold Spring Harb. Perspect. Med.* **6**, a026435 (2016).
3. Clarke, S. G. Protein methylation at the surface and buried deep: thinking outside the histone box. *Trends Biochem. Sci.* **38**, 243–252 (2013).
4. Cao, X. J. & Garcia, B. A. Global proteomics analysis of protein lysine methylation. *Curr. Protoc. Protein Sci.* **86**, 24.8.1–24.8.19 (2016).
5. Ambler, R. P. & Rees, M. W. ε-*N*-Methyl-lysine in bacterial flagellar protein. *Nature* **184**, 56–57 (1959).
6. Murray, K. The occurrence of epsilon-*N*-methyl lysine in histones. *Biochemistry* **3**, 10–15 (1964).
7. Tan, M. et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146**, 1016–1028 (2011).
8. Rea, S. et al. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593–599 (2000).
9. Elgin, S. C. & Reuter, G. Position-effect variegation, heterochromatin formation, and gene silencing in *Drosophila*. *Cold Spring Harb. Perspect. Biol.* **5**, a017780 (2013).
10. Wilkinson, A. W. et al. SETD3 is an actin histidine methyltransferase that prevents primary dystocia. *Nature* **565**, 372–376 (2019).
11. Petrossian, T. C. & Clarke, S. G. Uncovering the human methyltransferasome. *Mol. Cell. Proteom.* **10**, M110.000976 (2011).
12. Kuo, A. J. et al. NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. *Mol. Cell* **44**, 609–620 (2011).
13. Edmunds, J. W., Mahadevan, L. C. & Clayton, A. L. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J.* **27**, 406–420 (2008).
14. Schotta, G. et al. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev.* **22**, 2048–2061 (2008).
15. Beck, D. B., Oda, H., Shen, S. S. & Reinberg, D. PR-Set7 and H4K20me1: at the crossroads of genome integrity, cell cycle, chromosome condensation, and transcription. *Genes Dev.* **26**, 325–337 (2012).
16. Kuo, A. J. et al. The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. *Nature* **484**, 115–119 (2012).
17. McKay, D. J. et al. Interrogating the function of metazoan histones using engineered gene clusters. *Dev. Cell* **32**, 373–386 (2015).
18. Kaniskan, H. U. & Jin, J. Recent progress in developing selective inhibitors of protein methyltransferases. *Curr. Opin. Chem. Biol.* **39**, 100–108 (2017).
19. Carlson, S. M. et al. A proteomic strategy identifies lysine methylation of splicing factor snRNP70 by the SETMAR enzyme. *J. Biol. Chem.* **290**, 12040–12047 (2015).
20. Mazur, P. K. et al. SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer. *Nature* **510**, 283–287 (2014).
21. Roqueta-Rivera, M. et al. SETDB2 links glucocorticoid to lipid metabolism through Insig2a regulation. *Cell Metab.* **24**, 474–484 (2016).
22. Mas-Y-Mas, S. et al. The human mixed lineage leukemia 5 (MLL5), a sequentially and structurally divergent SET domain-containing protein with no intrinsic catalytic activity. *PLoS One* **11**, e0165139 (2016).
23. Fujiki, R. et al. Retraction: GlcNAcylation of a histone methyltransferase in retinoic-acid-induced granulopoiesis. *Nature* **505**, 574 (2014).
24. Osipovich, A. B., Gangula, R., Vianna, P. G. & Magnuson, M. A. Setd5 is essential for mammalian development and the co-transcriptional regulation of histone acetylation. *Development* **143**, 4595–4607 (2016).
25. Deliu, E. et al. Haploinsufficiency of the intellectual disability gene SETD5 disturbs developmental gene expression and cognition. *Nat. Neurosci.* **21**, 1717–1727 (2018).
26. Fnu, S. et al. Methylation of histone H3 lysine 36 enhances DNA repair by nonhomologous end-joining. *Proc. Natl Acad. Sci. USA* **108**, 540–545 (2011).
27. Huang, J. et al. Repression of p53 activity by Smyd2-mediated methylation. *Nature* **444**, 629–632 (2006).
28. Tan, X., Rotllant, J., Li, H., De Deyne, P. & Du, S. J. SmyD1, a histone methyltransferase, is required for myofibril organization and muscle contraction in zebrafish embryos. *Proc. Natl Acad. Sci. USA* **103**, 2713–2718 (2006).
29. Stender, J. D. et al. Control of proinflammatory gene programs by regulated trimethylation and demethylation of histone H4K20. *Mol. Cell* **48**, 28–38 (2012).
30. Eom, G. H. et al. Histone methyltransferase SETD3 regulates muscle differentiation. *J. Biol. Chem.* **286**, 34733–34742 (2011).
31. Fog, C. K., Galli, G. G. & Lund, A. H. PRDM proteins: important players in differentiation and disease. *BioEssays* **34**, 50–60 (2012).
32. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).
33. Li, J., Ahn, J. H. & Wang, G. G. Understanding histone H3 lysine 36 methylation and its deregulation in disease. *Cell. Mol. Life Sci.* **76**, 2899–2916 (2019).
34. Jha, D. K., Pfister, S. X., Humphrey, T. C. & Strahl, B. D. SET-ting the stage for DNA repair. *Nat. Struct. Mol. Biol.* **21**, 655–657 (2014).
35. Guo, R. et al. BS69/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing. *Mol. Cell* **56**, 298–310 (2014).
36. Wen, H. et al. ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. *Nature* **508**, 263–268 (2014).

37. Baubec, T. et al. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* **520**, 243–247 (2015).
38. Blackledge, N. P. et al. CpG islands recruit a histone H3 lysine 36 demethylase. *Mol. Cell* **38**, 179–190 (2010).
39. Bennett, R. L., Swaroop, A., Troche, C. & Licht, J. D. The role of nuclear receptor–binding SET domain family histone lysine methyltransferases in cancer. *Cold Spring Harb. Perspect. Med.* **7**, a026708 (2017).
40. Duns, G. et al. Histone methyltransferase gene SETD2 is a novel tumor suppressor gene in clear cell renal cell carcinoma. *Cancer Res.* **70**, 4287–4291 (2010).
41. Dalgliesh, G. L. et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463**, 360–363 (2010).
42. Collisson, E. A. et al. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543–550 (2014).
43. Zhang, J. et al. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* **481**, 157–163 (2012).
44. Zhu, X. et al. Identification of functional cooperative mutations of SETD2 in human acute leukemia. *Nat. Genet.* **46**, 287–293 (2014).
45. Parker, H. et al. Genomic disruption of the histone methyltransferase SETD2 in chronic lymphocytic leukaemia. *Leukemia* **30**, 2179–2186 (2016).
46. Roberti, A. et al. Type II enteropathy-associated T-cell lymphoma features a unique genomic profile with highly recurrent SETD2 alterations. *Nat. Commun.* **7**, 12602 (2016).
47. McKinney, M. et al. The genetic basis of hepatosplenic T-cell lymphoma. *Cancer Discov.* **7**, 369–379 (2017).
48. Martinelli, G. et al. SETD2 and histone H3 lysine 36 methylation deficiency in advanced systemic mastocytosis. *Leukemia* **32**, 139–148 (2018).
49. Viaene, A. N. et al. SETD2 mutations in primary central nervous system tumors. *Acta Neuropathol. Commun.* **6**, 123 (2018).
50. Gui, Y. et al. Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder. *Nat. Genet.* **43**, 875–878 (2011).
51. Huang, K. K. et al. SETD2 histone modifier loss in aggressive GI stromal tumours. *Gut* **65**, 1960–1972 (2016).
52. Gerlinger, M. et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
53. Hakimi, A. A. et al. Adverse outcomes in clear cell renal cell carcinoma with mutations of 3p21 epigenetic regulators *BAP1* and *SETD2*: a report by MSKCC and the KIRC TCGA research network. *Clin. Cancer Res.* **19**, 3259–3267 (2013).
54. Singh, R. R. et al. Intratumoral morphologic and molecular heterogeneity of rhabdoid renal cell carcinoma: challenges for personalized therapy. *Mod. Pathol.* **28**, 1225–1235 (2015).
55. Mar, B. G. et al. Mutations in epigenetic regulators including SETD2 are gained during relapse in paediatric acute lymphoblastic leukaemia. *Nat. Commun.* **5**, 3469 (2014).
56. Lee, J. J.-K. et al. Tracing oncogene rearrangements in the mutational history of lung adenocarcinoma. *Cell* **177**, 1842–1857 (2019).
57. Berquam-Vrieze, K. E. et al. Cell of origin strongly influences genetic selection in a mouse model of T-ALL. *Blood* **118**, 4646–4656 (2011).
58. Bard-Chapeau, E. A. et al. Transposon mutagenesis identifies genes driving hepatocellular carcinoma in a chronic hepatitis B mouse model. *Nat. Genet.* **46**, 24–32 (2014).
59. March, H. N. et al. Insertional mutagenesis identifies multiple networks of cooperating genes driving intestinal tumorigenesis. *Nat. Genet.* **43**, 1202–1209 (2011).
60. Rogers, Z. N. et al. A quantitative and multiplexed approach to uncover the fitness landscape of tumor suppression in vivo. *Nat. Methods* **14**, 737–742 (2017).
61. Walter, D. M. et al. Systematic in vivo inactivation of chromatin-regulating enzymes identifies Setd2 as a potent tumor suppressor in lung adenocarcinoma. *Cancer Res.* **77**, 1719–1729 (2017).
62. Xu, Q. et al. SETD2 regulates the maternal epigenome, genomic imprinting and embryonic development. *Nat. Genet.* **51**, 844–856 (2019).
63. Li, F. et al. The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSα. *Cell* **153**, 590–600 (2013).
64. Park, I. Y. et al. Dual chromatin and cytoskeletal remodeling by SETD2. *Cell* **166**, 950–962 (2016).
65. Chen, K. et al. Methyltransferase SETD2-mediated methylation of STAT1 is critical for interferon antiviral activity. *Cell* **170**, 492–506.e14 (2017).
66. Li, Y. et al. The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate. *J. Biol. Chem.* **284**, 34283–34295 (2009).
67. Papillon-Cavanagh, S. et al. Impaired H3K36 methylation defines a subset of head and neck squamous cell carcinomas. *Nat. Genet.* **49**, 180–185 (2017).
68. Zhu, L. et al. ASH1L links histone H3 lysine 36 dimethylation to MLL leukemia. *Cancer Discov.* **6**, 770–783 (2016).
69. Anderson, K. C. & Carrasco, R. D. Pathogenesis of myeloma. *Annu. Rev. Pathol.* **6**, 249–274 (2011).
70. Chng, W. J., Glebov, O., Bergsagel, P. L. & Kuehl, W. M. Genetic events in the pathogenesis of multiple myeloma. *Best. Pract. Res. Clin. Haematol.* **20**, 571–596 (2007).
71. Palumbo, A. & Anderson, K. Multiple myeloma. *N. Engl. J. Med.* **364**, 1046–1060 (2011).
72. Keats, J. J. et al. In multiple myeloma, t(4;14)(p16; q32) is an adverse prognostic factor irrespective of FGFR3 expression. *Blood* **101**, 1520–1529 (2003).
73. Santra, M., Zhan, F., Tian, E., Barlogie, B. & Shaughnessy, J. Jr. A subset of multiple myeloma harboring the t(4;14)(p16;q32) translocation lacks FGFR3 expression but maintains an IGH/MMSET fusion transcript. *Blood* **101**, 2374–2376 (2003).
74. Chesi, M. et al. The t(4;14) translocation in myeloma dysregulates both FGFR3 and a novel gene, MMSET, resulting in IgH/MMSET hybrid transcripts. *Blood* **92**, 3025–3034 (1998).
75. Aytes, A. et al. NSD2 is a conserved driver of metastatic prostate cancer progression. *Nat. Commun.* **9**, 5201 (2018).
76. Martinez-Garcia, E. et al. The MMSET histone methyl transferase switches global histone methylation and alters gene expression in t(4;14) multiple myeloma cells. *Blood* **117**, 211–220 (2011).
77. Jaffe, J. D. et al. Global chromatin profiling reveals NSD2 mutations in pediatric acute lymphoblastic leukemia. *Nat. Genet.* **45**, 1386–1391 (2013).
78. Oyer, J. A. et al. Point mutation E1099K in MMSET/NSD2 enhances its methyltransferase activity and leads to altered global chromatin methylation in lymphoid malignancies. *Leukemia* **28**, 198–201 (2014).
79. Carroll, W.L. et al. Pediatric acute lymphoblastic leukemia. in *ASH Education: Hematology* **2003**, 102–131 https://doi.org/10.1182/asheducation-2003.1.102 (2003).
80. Huang, C. & Zhu, B. Roles of H3K36-specific histone methyltransferases in transcription: antagonizing silencing and safeguarding transcription fidelity. *Biophys. Rep.* **4**, 170–177 (2018).
81. Cerami, E. et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).
82. Sankaran, S. M. & Gozani, O. Characterization of H3.3K36M as a tool to study H3K36 methylation in cancer cells. *Epigenetics* **12**, 917–922 (2017).
83. Sankaran, S. M., Wilkinson, A. W., Elias, J. E. & Gozani, O. A PWWP domain of histone-lysine *N*-methyltransferase NSD2 binds to dimethylated Lys-36 of histone H3 and regulates NSD2 function at chromatin. *J. Biol. Chem.* **291**, 8465–8474 (2016).
84. Wang, G. G., Cai, L., Pasillas, M. P. & Kamps, M. P. NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. *Nat. Cell Biol.* **9**, 804–812 (2007).
85. Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **517**, 576–582 (2015).
86. Taketani, T. et al. NUP98-NSD3 fusion gene in radiation-associated myelodysplastic syndrome with t(8;11)(p11; p15) and expression pattern of NSD family genes. *Cancer Genet. Cytogenet.* **190**, 108–112 (2009).
87. Shen, C. et al. NSD3-short is an adaptor protein that couples BRD4 to the CHD8 chromatin remodeler. *Mol. Cell* **60**, 847–859 (2015).
88. Kurotaki, N. et al. Haploinsufficiency of NSD1 causes Sotos syndrome. *Nat. Genet.* **30**, 365–366 (2002).
89. Gibson, W. T. et al. Mutations in EZH2 cause Weaver syndrome. *Am. J. Hum. Genet.* **90**, 110–118 (2012).
90. Douglas, J. et al. NSD1 mutations are the major cause of Sotos syndrome and occur in some cases of Weaver syndrome but are rare in other overgrowth phenotypes. *Am. J. Hum. Genet.* **72**, 132–143 (2003).
91. Tlemsani, C. et al. *SETD2* and *DNMT3A* screen in the Sotos-like syndrome French cohort. *J. Med. Genet.* **53**, 743–751 (2016).
92. Rolando, M. et al. *Legionella pneumophila* effector RomA uniquely modifies host chromatin to repress gene expression and promote intracellular bacterial replication. *Cell Host Microbe* **13**, 395–405 (2013).
93. Metzger, E. et al. KMT9 monomethylates histone H4 lysine 12 and controls proliferation of prostate cancer cells. *Nat. Struct. Mol. Biol.* **26**, 361–371 (2019).
94. Reynoird, N. et al. Coordination of stress signals by the lysine methyltransferase SMYD2 promotes pancreatic cancer. *Genes Dev.* **30**, 772–785 (2016).
95. Hamamoto, R. et al. SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat. Cell Biol.* **6**, 731–740 (2004).
96. Kunizaki, M. et al. The lysine 831 of vascular endothelial growth factor receptor 1 is a novel target of methylation by SMYD3. *Cancer Res.* **67**, 10759–10765 (2007).
97. Pinheiro, I. et al. Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. *Cell* **150**, 948–960 (2012).
98. Zhou, B. et al. PRDM16 suppresses MLL1r leukemia via intrinsic histone methyltransferase activity. *Mol. Cell* **62**, 222–236 (2016).

99. Campaner, S. et al. The methyltransferase Set7/9 (Setd7) is dispensable for the p53-mediated DNA damage response in vivo. *Mol. Cell* **43**, 681–688 (2011).
100. Mihola, O., Trachtulec, Z., Vlcek, C., Schimenti, J. C. & Forejt, J. A mouse speciation gene encodes a meiotic histone H3 methyltransferase. *Science* **323**, 373–375 (2009).
101. Takata, A. et al. Loss-of-function variants in schizophrenia risk and SETD1A as a candidate susceptibility gene. *Neuron* **82**, 773–780 (2014).
102. Tusi, B. K. et al. Setd1a regulates progenitor B-cell-to-precursor B-cell development through histone H3 lysine 4 trimethylation and Ig heavy-chain rearrangement. *FASEB J.* **29**, 1505–1515 (2015).
103. Palumbo, O. et al. Microdeletion of 12q24.31: report of a girl with intellectual disability, stereotypies, seizures and facial dysmorphisms. *Am. J. Med. Genet. A.* **167A**, 438–444 (2015).
104. Schmidt, K. et al. The H3K4 methyltransferase Setd1b is essential for hematopoietic stem and progenitor cell homeostasis in mice. *eLife* **7**, e27157 (2018).
105. Jones, W. D. et al. De novo mutations in MLL cause Wiedemann-Steiner syndrome. *Am. J. Hum. Genet.* **91**, 358–364 (2012).
106. Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A. J. & Korsmeyer, S. J. Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378**, 505–508 (1995).
107. Glaser, S. et al. Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* **133**, 1423–1432 (2006).
108. Lee, J. et al. Targeted inactivation of MLL3 histone H3-Lys-4 methyltransferase activity in the mouse reveals vital roles for MLL3 in adipogenesis. *Proc. Natl Acad. Sci. USA* **105**, 19229–19234 (2008).
109. Lee, J.-E. et al. H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *eLife* **2**, e01503 (2013).
110. Zech, M. et al. Haploinsufficiency of KMT2B, encoding the lysine-specific histone methyltransferase 2B, results in early-onset generalized dystonia. *Am. J. Hum. Genet.* **99**, 1377–1387 (2016).
111. McMahon, K. A. et al. Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. *Cell Stem Cell* **1**, 338–345 (2007).
112. Kleefstra, T. et al. Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. *Am. J. Hum. Genet.* **91**, 73–82 (2012).
113. Ng, S. B. et al. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat. Genet.* **42**, 790–793 (2010).
114. Tachibana, M. et al. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev.* **16**, 1779–1791 (2002).
115. Schaefer, A. et al. Control of cognition and adaptive behavior by the GLP/G9a epigenetic suppressor complex. *Neuron* **64**, 678–691 (2009).
116. Kleefstra, T. et al. Loss-of-function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome. *Am. J. Hum. Genet.* **79**, 370–377 (2006).
117. Tachibana, M. et al. Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev.* **19**, 815–826 (2005).
118. Ohno, H., Shinoda, K., Ohyama, K., Sharp, L. Z. & Kajimura, S. EHMT1 controls brown adipose cell fate and thermogenesis through the PRDM16 complex. *Nature* **504**, 163–167 (2013).
119. Dodge, J. E., Kang, Y. K., Beppu, H., Lei, H. & Li, E. Histone H3-K9 methyltransferase ESET is essential for early development. *Mol. Cell. Biol.* **24**, 2478–2486 (2004).
120. Liu, S. et al. Setdb1 is required for germline development and silencing of H3K9me3-marked endogenous retroviruses in primordial germ cells. *Genes Dev.* **29**, 108 (2015).
121. Ezhkova, E. et al. EZH1 and EZH2 cogovern histone H3K27 trimethylation and are essential for hair follicle homeostasis and wound repair. *Genes Dev.* **25**, 485–498 (2011).
122. Hidalgo, I. et al. Ezh1 is required for hematopoietic stem cell maintenance and prevents senescence-like cell cycle arrest. *Cell Stem Cell* **11**, 649–662 (2012).
123. O'Carroll, D. et al. The polycomb-group gene Ezh2 is required for early mouse development. *Mol. Cell. Biol.* **21**, 4330–4336 (2001).
124. Rayasam, G. V. et al. NSD1 is essential for early post-implantation development and has a catalytically active SET domain. *EMBO J.* **22**, 3153–3163 (2003).
125. Baujat, G. et al. Paradoxical NSD1 mutations in Beckwith-Wiedemann syndrome and 11p15 anomalies in Sotos syndrome. *Am. J. Hum. Genet.* **74**, 715–720 (2004).
126. Wright, T. J. et al. A transcript map of the newly defined 165 kb Wolf-Hirschhorn syndrome critical region. *Hum. Mol. Genet.* **6**, 317–324 (1997).
127. Lozier, E. R. et al. De novo nonsense mutation in WHSC1 (NSD2) in patient with intellectual disability and dysmorphic features. *J. Hum. Genet.* **63**, 919–922 (2018).
128. Boczek, N. J. et al. Developmental delay and failure to thrive associated with a loss-of-function variant in WHSC1 (NSD2). *Am. J. Med. Genet. A.* **176**, 2798–2802 (2018).
129. Nimura, K. et al. A histone H3 lysine 36 trimethyltransferase links Nkx2-5 to Wolf-Hirschhorn syndrome. *Nature* **460**, 287–291 (2009).
130. Chen, J. et al. Methyltransferase Nsd2 ensures germinal center selection by promoting adhesive interactions between B cells and follicular dendritic cells. *Cell Rep.* **25**, 3393–3404.e6 (2018).
131. Okamoto, N. et al. Novel MCA/ID syndrome with ASH1L mutation. *Am. J. Med. Genet.* **173**, 1644–1648 (2017).
132. Zhu, T. et al. Histone methyltransferase Ash1L mediates activity-dependent repression of neurexin-1α. *Sci. Rep.* **6**, 26597 (2016).
133. Jih, G. et al. The Trithorax-group protein ASH1L regulates hematopoietic stem cell homeostasis independently of its histone methyltransferase activity. *Blood* **132**(Suppl. 1), 1270 (2018).
134. Luscan, A. et al. Mutations in SETD2 cause a novel overgrowth condition. *J. Med. Genet.* **51**, 512–517 (2014).
135. Lumish, H. S., Wynn, J., Devinsky, O. & Chung, W. K. SETD2 mutation in a child with autism, intellectual disabilities and epilepsy. *J. Autism Dev. Disord.* **45**, 3764–3770 (2015).
136. Hu, M. et al. Histone H3 lysine 36 methyltransferase Hypb/Setd2 is required for embryonic vascular remodeling. *Proc. Natl Acad. Sci. USA* **107**, 2956–2961 (2010).
137. Wang, L. et al. H3K36 trimethylation mediated by SETD2 regulates the fate of bone marrow mesenchymal stem cells. *PLoS Biol.* **16**, e2006522 (2018).
138. Yi, X. et al. Histone methyltransferase Setd2 is critical for the proliferation and differentiation of myoblasts. *Biochim. Biophys. Acta* **1864**, 697–707 (2017).
139. Zuo, X. et al. The histone methyltransferase SETD2 is required for expression of acrosin-binding protein 1 and protamines and essential for spermiogenesis in mice. *J. Biol. Chem.* **293**, 9188–9197 (2018).
140. Skucha, A. et al. MLL-fusion-driven leukemia requires SETD2 to safeguard genomic integrity. *Nat. Commun.* **9**, 1983 (2018).
141. Jones, B. et al. The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS Genet.* **4**, e1000190 (2008).
142. Jo, S. Y., Granowicz, E. M., Maillard, I., Thomas, D. & Hess, J. L. Requirement for Dot1l in murine postnatal hematopoiesis and leukemogenesis by MLL translocation. *Blood* **117**, 4759–4768 (2011).
143. Nguyen, A. T., He, J., Taranova, O. & Zhang, Y. Essential role of DOT1L in maintaining normal adult hematopoiesis. *Cell Res.* **21**, 1370–1373 (2011).
144. Oda, H. et al. Monomethylation of histone H4-lysine 20 is involved in chromosome structure and stability and is essential for mouse development. *Mol. Cell. Biol.* **29**, 2278–2295 (2009).
145. Faundes, V. et al. Histone lysine methylases and demethylases in the landscape of human developmental disorders. *Am. J. Hum. Genet.* **102**, 175–187 (2018).

Acknowledgements

This work was supported in part by grants from the US National Institutes of Health to O.G. (R01GM079641). D.H. is supported by T32 AG0047126.

Competing interests

O.G. is a cofounder of Epicypher, Inc., and Athelas Therapeutics, Inc.

Additional information

Correspondence should be addressed to O.G.

Peer review information Anke Sparmann was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.