Published in final edited form as:
Dev Biol. 2010 March 15; 339(2): 240–249. doi:10.1016/j.ydbio.2009.08.017.

Histone H3 Lysine 4 (H3K4) Methylation in Development and Differentiation

Joel C. Eissenberg[1,*] and **Ali Shilatifard**[2,*]

[1]Edward A. Doisy Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, 1100 South Grand Blvd. St. Louis, MO 63104

[2]Stowers Institute for Medical Research, 1000 East 50th Street, Kansas City, MO 64110

Abstract

Covalent modification of histones on chromatin is a dynamic mechanism by which various nuclear processes are regulated. Methylation of histone H3 on lysine 4 (H3K4) implemented by the macromolecular complex COMPASS and its related complexes is associated with transcriptionally active regions of chromatin. Enzymes that catalyze H3K4 methylation were initially characterized genetically as regulators of *Hox* loci, long before their catalytic functions were recognized. Since their discovery, genetic and biochemical studies of H3K4 methylases and demethylases have provided important mechanistic insight into the role of H3K4 methylation in *HOX* gene regulation during development.

Keywords

histone methylation; trithorax; SET domain; Compass Complex; MLL; *HOX* genes

Introduction

The linear sequence of nucleotides in DNA is the primary storage form of genetic information in all living cells and in viruses. Execution of the genetic program encoded in DNA, on the other hand, relies on elaborate combinations of protein-DNA interactions, and in some cases, covalent modifications of chromatin. In eukaryotic nuclei, the DNA is complexed with an equal mass of histone protein. The DNA is wrapped around the outside of an octamer of two each of histones H2A, H2B, H3 and H4 or some variant of these canonical histones (Kornberg, 1974; Kornberg and Lorch, 1999). The N-terminal tails of each of the histones, however, protrude beyond the gyres of the DNA (Luger et al., 1997) and present themselves as substrates for a variety of covalent modifications including phosphorylation, acetylation, ubiquitination, ADP-ribosylation, biotinylation and methylation. Both the amino acid sequence and the targets of modification are, for the most part, conserved among all eukaryotes examined. Since the discovery of histone modifications in the 1960s and 1970s, much evidence has accumulated pointing to roles for specific modifications in mechanisms of developmental gene regulation and DNA repair (Shilatifard, 2006).

While lysine methylation of histones was first described in 1964 (Murray, 1964), decades passed without providing much insight into the functional significance of this modification. Most or all of the thinking on the significance of lysine methylation was guided by the belief

*Correspondence: eissenjc@slu.edu (J.C. Eissenberg), ash@stowers.org (A. Shilatifard).

that the role of lysine methylation in histones might affect the conformation, hydrophobicity or DNA affinity of the methylated histone (Honda et al., 1975). And indeed, it was recently shown that methylation of H4 at lysine 20 enhanced the ability of nucleosomal arrays to fold and condense *in vitro*, demonstrating that histone methylation can affect higher order chromatin structure directly (Lu et al., 2008). However, it has also become apparent that lysine methylation can create a binding site for proteins that can alter the local properties of chromatin for transcription (Fischle et al., 2008; Kouzarides, 2007). In the case of H3K4 methylation, this mark is generally associated with active transcription. H3K4 dimethylation appears to be broadly associated with active and potentially active genes, while H3K4 trimethylation is primarily a mark associated with the start site of transcription (Bernstein et al., 2002; Krogan et al., 2003; Ng et al., 2003). The first H3K4 methylase to be identified was Set1, from the yeast *S. cerevisiae,* within a complex named COMPASS (Complex Proteins Associated with Set1) (Miller et al., 2001). COMPASS is capable of mono- di- and trimethylating H3K4 (Krogan et al., 2002; Roguev et al., 2001; Shilatifard, 2006; Wood et al., 2007). Set1/COMPASS-related enzymes in animals and plants have also been identified, both genetically and biochemically. In most cases, these enzymes have been implicated in fundamental gene regulatory pathways during development. In this review, we summarize the major lines of genetic and biochemical data that link H3K4 methylation to developmental gene regulation in animals, plants and fungi.

Developmental regulation of gene expression by the *trithorax* gene family

Trithorax (*trx*) was discovered as a spontaneous mutation that gave a partial transformation of halteres, a pair of lobe-like projections that lie behind the wings, into wings (Ingham and Whittle, 1980). Further characterization revealed leg transformations as well, and in a few cases, the appearance of a supernumerary pair of wing-like structures on the thoracic segment ahead of the normal wing position, resulting in three sets of wings (hence the name "trithorax"; Fig. 1). As noted by Ingham (1998), the mutation *tetraptera*, described by Astauroff (1930) and since lost, had a similar phenotype and mapped to a similar chromosomal location and thus was likely to have been the first *trx* allele.

The phenotypes of the original *trx* allele (trx^1) included transformations of posterior abdominal segments to the appearance of more anterior segments, which, together with the haltere-to-wing transformation, resembles loss-of-function mutations in the *Bithorax* complex, one of the two clusters of *HOX* genes in the *Drosophila melanogaster* genome. This implicated *trx* as a positive regulator of *HOX* gene function. Subsequent cloning and molecular analysis established that the TRX protein shares a common motif, the eponymous Su(var)3–9/Enhancer of zeste/trithorax (SET) domain, with other so-called epigenetic regulators. Most interestingly, trx^{Z11} is a hypomorphic allele that contains a glycine-to-serine substitution in the SET domain (Breen, 1999; Stassen et al., 1995) that interferes with histone binding (Katsani et al., 2001).

Antibodies to TRX are capable of immunoprecipitating H3K4 methyltransferase activity from embryonic nuclear extract (Czermin et al., 2002), implicating either TRX or some associated protein in H3K4 methylation *in vivo*. The Drosophila TRX SET domain has been shown to have H3K4 methylase activity *in vitro*, and a mutated version of the TRX SET-domain carrying the same substitution found in the trx^{Z11} allele shows little or no methylation activity towards histone H3 (Smith et al., 2004). Thus, the genetic and biochemical characterization of TRX point to a role in maintaining gene activity, at least in part, through the methylation of H3K4.

Antibodies to TRX localize to ca. 60 sites on salivary gland polytene chromosomes, including the *HOX* gene clusters at the *Antennapedia* and *Bithorax* gene complexes

(Chinwalla et al., 1995). The number of chromosomal sites that accumulate significant levels of trimethylated H3K4 (H3K4me3) is far greater than the number of TRX binding sites inferred from immunostaining with this antibody (Tenney et al., 2006), suggesting either that TRX binding may be transient at most target sites or that there are other H3K4 trimethylases in Drosophila, albeit with non-redundant genetic activities. Recent chromatin immunprecipitation studies, however, suggest that the relationship between TRX protein binding sites and H3K4me3 may be complicated. While antibodies to the C-terminal region of TRX containing the SET domain show that binding sites for this domain have only limited overlap with sites of H3K4me3, antibodies to an N-terminal region of TRX shows that this domain is found at nearly every site of H3K4me3 enrichment (Schuettengruber et al., 2009). The distinct distributions of the N- and C-terminal TRX domains are consistent with proteolytic cleavage of TRX analogous to that seen with MLL (Hsieh et al., 2003) (Fig. 2). Proteolytic cleavage has been demonstrated for TRX and one of the two MLL cleavage sites is conserved in the TRX protein sequence (Capotosti et al., 2007). This site is deleted in the trx^{E3} allele, which causes impaired expression of *Antennapedia* Complex *HOX* genes but not *Bithorax* Complex *HOX* genes (Breen, 1999; Mazo et al., 1990; Sedkov et al., 1994). Thus, the functional linkage between TRX methylation activity and specific sites of H3K4 trimethylation remains ambiguous.

SET1, COMPASS and H3K4 methylation in yeast

To better understand the molecular function of TRX, the yeast Set1 protein was identified as the TRX homologue in yeast (Nislow et al., 1997), and was subsequently purified in a macromolecular complex named COMPASS (Miller et al., 2001). It was demonstrated that COMPASS has H3K4 methylase activity (Krogan et al., 2002; Miller et al., 2001; Roguev et al., 2001; Shilatifard, 2006). Subsequently, it was demonstrated that the mammalian MLLs also exist in COMPASS-like complexes capable of methylating H3K4 (Hughes et al., 2004; Shilatifard, 2008). COMPASS-like complexes in Drosophila have been reported, and antibodies to TRX are capable of immunoprecipitating H3K4 methyltransferase activity from embryonic nuclear extract (Czermin et al., 2002), implicating either TRX or some associated protein in H3K4 methylation *in vivo*.

Trithorax-related and ecdysone-mediated gene expression in Drosophila

Trithorax-related (*trr*) was first identified as a cDNA clone containing a SET domain using a degenerate PCR strategy (Sedkov et al., 1999). A putative null allele, trr^{1}, is recessive embryonic lethal. No evidence for homeotic transformations resembling loss-of-function *HOX* gene mutations was observed in the dead embryos. Furthermore, no dominant interaction was seen between trr^{1} and hypomorphic alleles of either *Polycomb* or *trithorax*, suggesting that *trr* is probably not a major regulator of *HOX* genes.

Trr encodes an H3K4 trimethylase (Sedkov et al., 2003). The TRR protein is more abundant in embryos than TRX or the H3K4 methylase ASH1 (discussed below) and more widely distributed than these other proteins on polytene chromosomes. TRR co-localizes extensively with the ecdysone receptor (EcR), the receptor for the major steroid hormone in Drosophila. Antibody to the EcR can co-precipitate TRR in an ecdysone-dependent fashion from embryo extract, suggesting that TRR and the EcR are part of a ligand-dependent common complex. Chromatin immunoprecipitation in ecdysone-treated S2 cultured cells shows that two ecdysone-induced promoters are associated with elevated levels of EcR, TRR and trimethylated H3K4. Importantly, transcription and H3K4 trimethylation are markedly reduced at both loci in embryos homozygous for the trr^{1} allele compared to wild-type embryos. A similar reduction is observed in embryos homozygous for the trr^{3} allele, which deletes the SET domain of TRR, implicating the catalytic activity of TRR in the

mechanism of methylation and ecdysone-dependent gene activation *in vivo*. Thus, although they share many structural similarities and catalytic activities, TRX and TRR have diverged in their developmental roles.

The mammalian trx gene family

A similar specialization appears to have occurred among the TRX family proteins in mammals, suggesting an ancient origin for this division of labor (Goo et al., 2003; Lee et al., 2008).The mammalian homologs of *trx* include the MLL (Mixed Lineage Leukemia) protein (Shilatifard, 2008). MLL and its translocations are associated with the pathogenesis of hematological malignancies (Krivtsov and Armstrong, 2007; Rowley, 1998; Tenney and Shilatifard, 2005). The mammalian genome contains genes encoding six *trx*-related proteins: Set1A and Set1B, MLL, MLL2, MLL3, and MLL4. All are found in COMPASS-like complexes capable of methylating H3K4 (Fig. 3) (Cho et al., 2007; Hughes et al., 2004; Lee and Skalnik, 2005; Lee et al., 2007a; Shilatifard, 2008; Wu et al., 2008). There is only one H3K4 methylase in yeast, and it is not clear why mammals carry at least six, functionally non-redundant H3K4 HMTases.

The *Mll* gene (also called ALL-1 and HRX) was discovered by virtue of its involvement in human acute leukemia; chromosome rearrangements involving a breakpoint within the MLL gene are associated with a variety of acute myeloid and lymphoid leukemias (Djabali et al., 1992; Gu et al., 1992; Mbangkollo et al., 1995; Tkachuk et al., 1992). While MLL is structurally homologous to Drosophila TRX, it also contains three AT-hook domains and a DNA methyltransferase homology domain that are absent from TRX.

Considerable experimental evidence points to a key role of MLL in hematopoiesis (reviewed in Dorshkind and Witte, 2004; Ono et al., 2005). Mice heterozygous for an *Mll1* knockout allele are anemic and have reduced platelet and B-cell counts (Yu et al., 1995). Homozygous *Mll1* knockout mouse embryos developed from chimeras of $Mll1^{+}$ and $Mll1^{-}$ embryonic stem cells still support the development of early hematopoietic progenitor cells (Ernst et al., 2004). However, these cells show proliferation and differentiation defects. Remarkably, ectopic *HOX* gene expression rescues the growth defect, strongly implicating *Mll1* in the maintenance of HOX genes required for hematopoiesis. Targeted knockout of *Mll1* in mouse hematopoietic stem cells implicates MLL in hematopoietic stem cell renewal (McMahon et al., 2007).

MLL and HOX gene regulation in mammals

MLL is essential and required for axial segment identity in mice (Yu et al., 1995). Mice that are heterozygous for an *Mll1* knockout allele show rostral skeletal transformations in the cervical, thoracic and lumbar vertebrae resembling the effects loss-of-function alleles for multiple HOX genes. Importantly, the initial axial patterns of *HOX* gene expression are established during embryogenesis, but are lost in heterozygous *Mll1* knockout animals, arguing for a maintenance role in HOX gene activation analogous to that of *trx* in Drosophila (Yu et al., 1998).

Although *Mll1* null mutations result in embryonic lethality, mice homozygous for an *Mll1* mutation lacking the SET domain (*Mll1ΔSET*) are viable and fertile (Terranova et al., 2006), demonstrating that MLL has essential functions distinct from H3K4 methylation. Nevertheless, *Mll1ΔSET* mice show skeletal defects resembling knockouts of specific *HOX* genes as well as reduced mono- and di-methylated H3K4 levels at the promoters of specific HOX gene promoters in the embryonic trunk tissue.

Chromatin immunoprecipitation analysis of MLL distribution in human monocyte and lymphoblast cell lines using a microarray consisting of human proximal promoter and

noncoding sequences found MLL at ca. 40% of promoters on the array, including ca. 90% of sites containing trimethylated H3K4 and occupied by RNA Polymerase II (Guenther et al., 2005). Using tiling arrays for a subset of *HOX* and non-*HOX* genes, both MLL and trimethylated H3K4 were found concentrated in promoter-proximal chromatin of highly expressed genes, as is the case with yeast Set1 (see below). Interestingly, the distribution of MLL and trimethyl H3K4 extends broadly upstream and downstream of the promoters of several late HOX genes, suggesting that MLL acts by different mechanisms at HOX and non-HOX loci, and could also maintain large domains of active chromatin specifically at HOX gene clusters. However, recent studies have demonstrated that less than 2% of the genes in mouse embryo fibroblast are the sole H3K4 methylation targets of MLL (Wang, Submitted).

Since all leukemia-associated MLL translocations involve the expression of a MLL fusion protein that deletes the C-terminal SET domain (Krivtsov and Armstrong, 2007; Tenney and Shilatifard, 2005), the relationship between the oncogenic activity of MLL fusion proteins and H3K4 methylation is unclear. Furthermore, the interpretation of the genetic and chromatin immunoprecipitation data concerning MLL is complicated by the fact that MLL is proteolytically cleaved *in vivo*, resulting in N- and C-terminal fragments that can associate non-covalently (Hsieh et al., 2003). As suggested by the chromatin immunoprecipitation data using antibodies to N- and C-terminal domains of TRX in Drosophila embryos (Schuettengruber et al., 2009), the distributions of the MLL N- and C-terminal fragments may be distinct *in vivo*. Similarly, the functions of MLL protein complexes containing only one fragment versus those containing both could be distinct.

MLL2 was cloned based on sequence homology to MLL (FitzGerald and Diaz, 1999; Huntsman et al., 1999). MLL2 was found within regions amplified in a subset of pancreatic and glioblastoma cell lines, although no mechanistic relationship between MLL2 amplification and these cancers has been established. Like MLL, MLL2 inactivation is embryonic lethal in mice (Glaser et al., 2006), demonstrating that these structurally homologous proteins are not functionally redundant. MLL and MLL2 are most highly similar to TRX in Drosophila and both have been shown to be able to regulate Hox gene expression in mammalian cells (Ernst et al., 2004; Glaser et al., 2006). Furthermore, both MLL and MLL2 are found in similar complexes, both containing menin (Fig. 3).

Human TRX homologs and nuclear hormone receptors

MLL4/ALR was identified as a cDNA clone in a low stringency hybridization screen using the MLL SET domain coding sequence as a probe (Prasad et al., 1997). Similarly, the genes that require MLL4/ALR appear to be distinct (Prasad et al., 1997). A complex containing MLL4/ALR was found to interact with the estrogen receptor (Mo et al., 2006). siRNA to MLL4/ALR reduced receptor-dependent transcription, implicating MLL4/ALR as a co-activator in receptor-induced transcription. In this respect, MLL4/ALR seems to be a functional homolog of TRR in Drosophila--a co-activator of genes controlled by steroid hormone receptors.

MLL3 was identified based on sequence similarity to the other MLL-related proteins during sequence analysis of human chromosome 7 (Ruault et al., 2002) and independently in a yeast two-hybrid protein screen of a human brain cDNA library for interactors with the nuclear protein HUEL (Tan and Chow, 2001). It maps to a region of chromosome 7 that is frequently deleted in patients with developmental defects and leukemia.

MLL3 and MLL4/ALR have been found in association with retinoic acid receptor, implicating them in nuclear receptor-mediated gene activation (Goo et al., 2003). RNAi against both MLL3 and MLL4/ALR causes reduced expression and H3K4 trimethylation of

a retinoic acid receptor target gene (Lee et al., 2006). Again, this would seem to place MLL3 and MLL4/ALR, in the category of nuclear receptor co-activators like Drosophila TRR. Indeed, the closest mammalian homologues of Drosophila TRR are the MLL3 and MLL4/ALR proteins. Given that mammals have considerably more nuclear receptor superfamily proteins than Drosophila, it may be that the MLL3–MLL4/ALR proteins may both have evolved from an ancestral TRR-like protein to accommodate this diversification

The studies to date clearly demonstrate that, despite their structural similarities, the MLL family of proteins in mammals have distinct and non-redundant functions in development. Defining the specific gene targets of each MLL protein and the mechanisms of targeting will be essential in understanding this important family of developmental regulators.

MLL5 and hematopoiesis in mammals

Mll5 was cloned as a gene with limited homology to MLL found in a region that is often deleted in patients with acute myeloid leukemia and myelodysplasitc syndrome who have poor clinical outcome (Emerling et al., 2002; Kratz et al., 2001; Le Beau et al., 1996; Luna-Fineman et al., 1995). However, no MLL5 mutations have been identified among primary leukemia cells examined (Emerling et al., 2002).

MLL5 is only about half the size of MLL, lacks the AT-hook, DNA methyltransferase, and FYRN-FYRC homology domains of MLL and its SET domain is near the center of the protein rather than near the C-terminus. Furthermore, based on sequence analysis, MLL5 is no more closely related to MLL than to other SET domain families such as the Set2 family, some of whose members methylate H3 at lysine 36. Furthermore, when shared components of the MLL1–4 and Set1A/Complexes such as ASH2, WDR5 or RBBP5 are tagged and purified, MLL5 is not identified in such purifications (Wu et al., 2008) (Fig. 3). Putative orthologs of MLL5 in other species include the largely uncharacterized CG9007 in Drosophila and SET3 and SET4 in yeast.

Despite the low similarity between MLL5 and Set1/MLL1–4, a complex containing MLL5 is capable of methylating nucleosomal histone H3, and peptide methylation analysis demonstrated a specificity for H3K4 (Fujiki et al., 2009). *Mll5* is expressed in all embryonic and adult tissues, including hematopoeitic organs (Emerling et al., 2002). Despite the high conservation level of MLL5 in different species, it is not essential for embryogenesis in mammals. However, mice in which MLL5 is inactivated have impaired immune response and show defects in several aspects of hematopoiesis (Heuser et al., 2009; Madan et al., 2009; Zhang et al., 2009).

Recently, MLL5 was found to be associated with retinoic acid receptor α(RAR α) in retinoic acid-treated granulocyte-like HL60 cells (Fujiki et al., 2009). Over-expressed MLL5 stimulates expression of an RAR α-dependent reporter in an SET domain-dependent fashion. MLL5 activity is potentiated by O-linked glycosylation at threonine 440. Interestingly, this threonine is conserved in the Drosophila MLL5 homolog, but not in yeast SET3 and SET4. Immunoprecipitated MLL5 complexes containing a T440A mutation in MLL5 are catalytically dead. Importantly, overexpression of wild-type MLL5 promotes the RAR α-dependent differentiation of cultured HL60 cells, while overexpression of either the SET domain-deleted MLL5 or the T440A mutation attenuated the response. Thus, MLL5 is implicated in RAR α-dependent H3K4 methylation, with an activity that is regulated by O-linked glycosylation. It would be interesting to determine whether developmental defects in mice carrying the T440A allele are as extensive as in the *Mll5* knockout.

Ash1 and homeotic gene regulation in Drosophila

Absent, small or homeotic discs1 (*ash1*) was first identified in a genetic screen for late larval lethal mutations that displayed abnormal or missing imaginal discs (Shearn et al., 1987; Shearn et al., 1971). Although most allelic combinations of Ash1 mutations are lethal, surviving and pharate adults display transformations including haltere to wing, first and third legs to second leg and genitalia to leg or antenna (Shearn, 1980; Shearn, 1989). These all represent anterior-ward transformations, similar to loss-of-function alleles of *Bithorax* complex genes, placing *ash1* in the *trithorax* group of homeotic regulators.

Using an *ash1* allelic series of increasing phenotypic severity, Byrd and Shearn (2003) found a progressive loss of chromosomal dimethyl H3K4, with a virtually complete loss of chromosomal dimethyl H3K4 with the strongest allelic combination. The ASH1 protein is an H3K4 methylase *in vitro*, (Byrd and Shearn, 2003), although it can also show methylation activity towards H3K9 and H4K20 (Beisel et al., 2002). However, larvae carrying strong mutant alleles of Ash1 show only a partial reduction of dimethylated H3K9 on polytene chromosomes and no detectable reduction in chromosomal H4K20 (Byrd and Shearn, 2003), suggesting the H3K4 is the primary substrate for ASH1 *in vivo*.

Genetic evidence supports a direct role for the ASH1 SET domain in H3K4 methylation by Ash1. The hypomorphic $Ash1^{10}$ allele carries an asparagine-to-isoleucine substitution in the SET domain that leads to dramatic reduction of H3K4 dimethylation when heterozygous with the null allele $Ash1^{22}$ (Byrd and Shearn, 2003). Furthermore, immunolocalization of ASH1 on larval salivary gland chromosomes shows that ASH1 protein is found at over 100 sites and has a much wider chromosomal distribution than TRX (Tripoulas et al., 1996). The mutation in the $Ash1^{10}$ allele results in somewhat reduced immunostaining at specific sites. About a quarter of the ASH1 sites overlap those of TRX, including the *Antp* HOX complex but not the *Bithorax* HOX complex (Rozovskaia et al., 1999; Tripoulas et al., 1996). However, ASH1 can be found by chromatin immunoprecipitation at several sites within the *Ubx* locus of the *Bithorax* complex in imaginal discs of the metathoracic legs and halteres (where *Ubx* is expressed) but not in wing discs and S2 cultured cells (where *Ubx* is silent) (Sanchez-Elsner et al., 2006). This suggests a tissue-dependent distribution for ASH1. The number of ASH1 sites is far larger than the number of homeotic genes, suggesting that ASH1 probably regulates many other targets during development.

ASH1 and TRX can be co-immunoprecipitated from embryonic extracts, and *in vitro* binding experiments suggest that they can interact through their SET domains (Rozovskaia et al., 1999). Given the dominant role of ASH1 in H3K4 dimethylation, one possible role for the TRX-ASH1 association could be for ASH1 to create a dimethyl-H3K4 substrate that can subsequently be trimethylated by TRX.

The human homolog of Ash1

Human Ash 1 (huAsh1) was first identified as an EST clone homologous to the Drosophila ASH1 SET domain (Nakamura et al., 2000). The SET domain of huAsh1 methylates H3K4 *in vitro* (Gregory et al., 2007). In contrast to the reported *in vitro* substrates of Drosophila ASH1, huAsh1 does not seem to methylate H4K20 or H3K9; indeed, H3K9 methylation impairs *in vitro* methylation of H3K4 by the huAsh1.

Chromatin immunoprecipitation finds huAsh1, MLL, and H3K4 trimethylation are co-enriched in the transcribed regions of active housekeeping genes, but not inactive genes. This association was independent of the phosphorylated isoform of RNA Polymerase II, the elongating form of the enzyme (Gregory et al., 2007), as it is for ASH1 in Drosophila (Eissenberg et al., 2007b). Significantly, both huAsh1 and MLL are enriched throughout the transcribed region of HOXA10 in HEK-293T cells, which express this HOX gene (Gregory

et al., 2007). The functional contribution of huAsh1 binding to transcription and H3K4 methylation was not unambiguously determined in this study due to incomplete knockdown of huASH1 protein. However, the functional interdependence between huAsh1 binding and MLL could be assessed in *Mll1* null mouse embryonic fibroblasts. For non-*HOX* genes examined, huAsh1 was still present at the 5' end of the genes; these genes showed some reduction in di- and tri-methylated H3K4 but normal transcription. Interestingly, in these cells the level of H3K4 trimethylation at HoxA9 is dramatically reduced, but levels of huAsh1 and H3K4 dimethylation are only moderately reduced. This is consistent with the finding that Drosophila ASH1 has a major role in H3K4 dimethylation (Byrd and Shearn, 2003).

Trithorax homologs and development in Arabidopsis

The *Arabadopsis thaliana* genome encodes nine *trithorax* family proteins (Alvarez-Venegas and Avramova, 2002; Baumbusch et al., 2001; Ng et al., 2007). The Arabidopsis homologue of TRX, ATX1, is required for floral development, and its SET domain has weak H3K4 methylating activity *in vitro* (Alvarez-Venegas et al., 2003). In plants homozygous for a loss-of-function ATX1 allele, defects in flower structures as well as transformations in flower organ identity are observed. The Arabidopsis genome encodes structural homologs of the mammalian COMPASS subunits WDR5, ASH2L and RBBP5, suggesting that the COMPASS complex is conserved in plants..

Plants don't have *HOX* genes, however, the MADS box family of transcription factors control flower organ identity (Honma and Goto, 2001; Pelaz et al., 2000; Riechmann and Meyerowitz, 1997). Reduced levels of steady state transcripts for several MADS box genes are observed in the developing buds of ATX1 mutant plants, consistent with a role for ATX1 in activating these genes in early flower development (Alvarez-Venegas et al., 2003).

Histone H2B monoubiquitination and H3K4 methylation crosstalk

Since methylation of the H3K4 within chromatin by yeast COMPASS is required for the proper regulation of gene expression, we set out to define the pathway controlling H3K4 modifications in yeast. We devised a method to screen the entire yeast proteome for proteins involved in this pathway termed Global Proteomic analysis in *S. cerevisiae* (GPS) (Schneider et al., 2004). Using GPS, it was demonstrated that ubiquitination of lysine 123 of histone H2B by the Rad6/Bre1 complex is required for histone methylation by COMPASS (Dover et al., 2002; Shilatifard, 2006; Wood et al., 2003a; Wood et al., 2003b) (Fig. 4). This process, termed histone crosstalk, is highly conserved from yeast to human (Pavri et al., 2006; Shilatifard, 2006; Smith and Shilatifard, 2009). This crosstalk is mediated through the Cps35 (Wdr82 in humans) subunit of COMPASS (Lee et al. 2007, Wu et al. 2008, Fig. 4). Due to the absence of Wdr82 in the MLL1–4 complexes, we have speculated that Set1, but not MLL1–4, works via this cross-talk pathway, but this idea awaits experimental validation (Smith and Shilatifard, 2009).

In addition to Rad6 and Bre1, GPS uncovered a role for the components of the Paf1 complex, a Pol II elongation factor, as well as the Pol II CTD kinase, and Bur1/Bur2 in the regulation of histone monoubiquitination and methylation (Krogan et al., 2003; Wood et al., 2003b; Wood et al., 2005) (Fig. 4). GPS was also instrumental in defining a role for another CTD kinase, Ctk1, in proper H3K4 mono-, di-, and trimethylation (Wood et al., 2005; Wood et al., 2007).

Studies from different laboratories during the past several years have concentrated on defining the pattern of histone modifications throughout the genome. We know that transcription-coupled activation of the H3K4 methylation mark can result in erasure of the

repressive H3K27 methylation mark. While the field has produced a wealth of information regarding histone marks, few details are known of the machinery and mechanisms involved in their creation and removal. In the next section, we will discuss recent findings regarding the molecular machinery required for the removal of H3K4 methylation.

H3K4 demethylases and development

In addition to the identification of factors required for proper H3K4 methylation, we and others have also identified the machinery required for demethylating H3K4, which is also conserved from yeast to mammals. The methylation of histones was long considered to be a biochemically stable mark, making it an attractive candidate for a mechanism of transcriptional memory. Metabolic labeling of cultured Chinese hamster ovary cells, developing rat brain, developing trout testes or cultured alfalfa cells indicated that little or no turnover of methyllysine occurs outside of protein turnover (Byvoet et al., 1972; Duerre and Lee, 1974; Honda et al., 1975; Waterborg, 1993), leading to the assumption that this modification is irreversible. Recent biochemical and genetic studies have uncovered several classes of histone demethylases (Agger et al., 2008; Shi and Whetstine, 2007; Smith et al., 2008; Trojer and Reinberg, 2006). Among them are lysine demethylases capable of reversing the H3K4 methyl mark.

Little imaginal discs (Lid) and gene regulation in Drosophila

Lid was first identified in a genetic screen for intergenic noncomplementation with *Ash1* mutations (Gildea et al., 2000). This phenotype implies that *Lid* belongs to the *trithorax* group of homeotic gene activators. This phenotype seems paradoxical in light of the subsequent discovery that *Lid* encodes a jumonji C domain-containing protein with trimethyl H3K4 demethylating activity (Eissenberg et al., 2007a; Lee et al., 2007b; Secombe et al., 2007).

How could a H3K4 demethylase be a TRX-like transcriptional activator if its role *in vivo* is to reverse the TRX/Ash1/TRR methyl mark? One possibility is that loss of H3K4 demethylation activity results in ectopic H3K4 methylation and the mislocalization of factors that bind the H3K4 methyl mark. One such factor is the chromo helicase CHD1, which has been shown to bind methylated H3K4 (Flanagan et al., 2005; Sims et al., 2005). In Drosophila, CHD1 is concentrated at transcriptionally active loci, requires its chromo domains in order to bind chromatin, and is lost from chromosomes under conditions where H3K4 trimethylation is reduced (Eissenberg et al., 2007b; Kelley et al., 1999; Stokes et al., 1996). Recent evidence implicates CHD1 in chromatin remodeling *in vivo* (Konev et al., 2007), remodeling that could facilitate gene activation. Thus, ectopic H3K4 methylation would be expected to titrate CHD1 away from its normal transcriptional targets, reducing expression of those genes. Indeed, Eissenberg et al. (2007a) showed that *Lid* knockdown in salivary gland chromosomes results in elevated genome-wide H3K4 trimethylation and a relative reduction of CHD1 at loci where it is normally concentrated.

Mutations in *Lid* have no homeotic phenotypes, although they can enhance the phenotype of a *Polycomb* hypomorphic allele (Gildea et al., 2000). Instead, hypomorphic alleles show a high frequency of duplicated thoracic bristles, suggesting a defect in the *Notch* signaling pathway. This suggests that either another H3K4 demethylase exists in Drosophila, or else that no enzymatic demethylase exists to oppose the TRX H3K4 methylation activity at HOX genes.

Role of the H3K4 mark in development

Does the H3K4 methyl mark constitute a form of molecular "memory" for the transcriptionally activated state? The answer to this question depends on what one means by

memory. In one sense, molecular memory is represented by the perdurance of the methyl mark after the enzyme has completed its catalysis. In that sense, the persistence of H3K4 di- and trimethylation at an induced gene for over an hour after gene inactivation and Set1 dissociation could constitute a kind of memory of the active state (Krogan et al., 2003; Ng et al., 2003). However, this mark doesn't seem to persist through mitosis, and there is no evidence that the elevated methylation results in a distinct mechanism of induction.

In developmental terms, can H3K4 methylation promote transcriptional activity if the factor required to establish the mark is removed? Using X-rays to induce somatic recombination at different periods of development and then inspecting adult cuticular markers, Ingham (1985) found that *trx* is required continuously through third instar to maintain segment identity. A later study confirmed that *trx*$^-$ clones in the leg lead to loss of UBX protein expression 96 hours later, consistent with a continuous requirement for TRX (Klymenko and Muller, 2004). If TRX is directly responsible for establishing the H3K4 methyl mark at specific HOX loci, this would suggest that the TRX-dependent H3K4 methyl mark alone is insufficient to maintain a "memory" of the *trx* activation pattern. Since this study didn't directly measure the H3K4 mark, it is unknown whether H3K4 methylation is lost when *trx* is removed.

A similar requirement for continuous activity for *Ash1* has been demonstrated in the activation of *Ubx* (Klymenko and Muller, 2004). Again, the status of H3K4 methylation at *Ubx* was not assessed, although the results of Byrd and Shearn (2003) would predict a complete loss of dimethylated H3K4. Analogous experiments have not been done with *Trr*.

This contrasts with another SET domain protein, SU(VAR)3–9, which methylates H3K9 and which has dosage-dependent effects on heterochromatic gene silencing. Heterochromatin silencing in the fly established by a single dose of the SET domain protein SU(VAR)3–9 is not reversed in clones of two-dose cells induced after embryogenesis. Thus, the "memory" of the embryonic state of suppressed silencing was still retained in the daughter cells in which SU(VAR)3–9 dose is restored to wild type levels (Rudolph et al., 2007). Here, too, the study didn't directly measure H3K9 methylation, so a role for the histone methyl mark in the "memory" mechanism is only inferred.

POLYCOMB-mediated silencing provides another test of molecular memory. In this case, *HOX* genes in Drosophila that are initially silenced by the action of segmentation genes are maintained in a silent state after the repressors that established the silent state have disappeared. In Drosophila, this silencing mechanism depends on the H3K27 methylase *Enhancer of zeste* [*E(z)*], a subunit of the E(Z)-ESC *Polycomb* repression complex. When null clones for *E(z)* are generated in the primordial of the wing or leg, ectopic *Ubx* expression appears in the *E(z)* null cells (Klymenko and Muller, 2004), demonstrating a requirement for E(Z) in the maintenance of *Ubx* repression.

The H3K4 methylases TRX and ASH1 are thought to antagonize the silencing of HOX genes by the POLYCOMB-containing complexes. Surprisingly, however, clones that are simultaneously lacking ASH1 and E(Z) still show ectopic *Ubx* expression. A similar ectopic *Ubx* expression is observed in cells doubly mutant for *trx* and *Sex combs on midleg* (*Scm*), a subunit of the distinct PRC1 *Polycomb* repression complex. Taken together, the results suggest that the TRX and ASH1 H3K4 methylases act as anti-repressors and not as transcriptional co-activators. Since the anti-repressor effect is lost after the developmental period at which HOX gene expression patterns are initially established, any "memory" of TRX/ASH1 activity must have been erased in the hours after the cells lost the enzymes.

Thus, on the timescale of development, the continuous presence of H3K4 methylases appears to be required to maintain their effects of promoting transcription. The role of the

H3K4 methyl mark is likely to be as a binding site for proteins that facilitate access by RNA Polymerase II to the DNA. In that sense, the memory conferred by the H3K4 methyl mark is transient, highly coupled to the presence of the methylase and rapidly reversed.

Concluding Remarks

H3K4 methylation and machinery involved in its implementation are evolutionarily conserved marks of active and potentially active genes in all eukaryotes examined (Shilatifard, 2006; Shilatifard, 2008). Remarkably, a role for the *trx* family of H3K4 methylases in the specification of tissue identity has been conserved in plants and animals, despite the fundamental differences in the mechanisms of homeosis in these kingdoms.

A model for the mechanism by which TRX family of proteins antagonize transcriptional silencing is that TRX acts directly to interfere with methylation of H3K9 and/or H3K27, histone marks recognized and bound by the HP1 and POLYCOMB chromo domain proteins, respectively. However, it is probably important to keep in mind that there may also exist nonhistone targets for methylation by any of these proteins, and that such substrates may be equally or more important than histones in the biological activities of at least some of these enzymes for at least some genes. Indeed, the HMTase homolog Set7 has been shown to methylate the transcription factors p53 and TAF10 (Chuikov et al., 2004; Kouskouti et al., 2004). The extent to which H3K4 methylases have nonhistone substrates is currently unknown.

While the correlation between H3K4 methylation (especially H3K4 trimethylation) and transcriptional activity is high in all eukaryotes examined, there are striking exceptions. Increasing numbers of examples of "bivalent" marks have been described, in which H3K4 methylation is enriched at loci that are also enriched for silencing-associated marks such as H3K9 or H3K27 methylation (Bernstein et al., 2006; Saleh et al., 2007). The significance of bivalently marked loci is unknown, although it is speculated that this chromatin state could keep a silent locus "poised" for activation, perhaps by recruiting both activators and silencers to the same domain. It is thought that such a bivalent state could underlie the pluripotency of stem cells, although the requirement for such bivalency in stem cell maintenance is untested.

There is no evidence that the H3K4 mark is stably maintained in the absence of methylase activity in proliferating cells. Thus, while TRX can be said to propagate a kind of epigenetic "memory" for the activity of *HOX* genes through development, and this activity of TRX does require its H3K4 methylase activity, there is no evidence that the methyl mark can maintain the transcriptional activation function of TRX in the absence of TRX. One confounding variable in uncoupling the role of H3K4 methylation from TRX is the role of cell division. With each round of DNA replication, new unmethylated H3 is deposited on the daughter strands and the marked nucleosomes will gradually be lost through dilution. A test of whether H3K4 methylation is itself a memory mark would be to test whether knockdown of TRX in nonproliferating cells results in loss of expression of a *trx*-dependent reporter gene.

Regardless of the mechanism by which the H3K4 methyl mark contributes to key regulatory and developmental fate decisions, the striking evolutionary conservation of this mark argues for an ancient origin. The number of enzymes capable of methylating H3K4 have proliferated with the split between plants and animals and with the radiation of metazoa, suggesting a diversification in the targeting of chromatin domains for H3K4 methylation during development and differentiation. We anticipate that the dissection of mechanisms for H3K4 targeting, as well as the mechanisms by with H3K4 methylation results in changes in

gene expression, will yield fundamental insights into the molecular circuitry of development and cell fate decisions.

Acknowledgments

We thank Dr. Edwin Smith for helpful comments on the manuscript and Dr. Zoya Avramova for helpful discussion on Arabidopsis COMPASS. We thank Dr. Philip Ingham and the International Journal of Developmental Biology for permission to use the image shown in Fig. 1B. The work in authors' laboratories is supported by NSF grant MCB 0724501 to JCE and NIH R01CA89455, R01GM069905 to AS.

References

Agger K, Christensen J, Cloos PA, Helin K. The emerging functions of histone demethylases. Curr Opin Genet Dev. 2008; 18:159–168. [PubMed: 18281209]

Agger K, Christensen J, Cloos PA, Helin K. The emerging functions of histone demethylases. Curr Opin Genet Dev. 2008; 18:159–168. [PubMed: 18281209]

Alvarez-Venegas R, Avramova Z. SET-domain proteins of the Su(var)3–9, E(z) and trithorax families. Gene. 2002; 285:25–37. [PubMed: 12039029]

Alvarez-Venegas R, Pien S, Sadder M, Witmer X, Grossniklaus U, Avramova Z. ATX-1, an Arabidopsis homolog of trithorax, activates flower homeotic genes. Curr Biol. 2003; 13:627–637. [PubMed: 12699618]

Astauroff BL. Analyse der erblichen Störungsfälle der bilateralen Symmetrie im Zusammenhang mit der selbständigen Variabilität ähnlicher Strukturen. Z. indukt. Abstamm. Vererb. 1930; 55:183–262.

Baumbusch LO, Thorstensen T, Krauss V, Fischer A, Naumann K, Assalkhou R, Schulz I, Reuter G, Aalen RB. The Arabidopsis thaliana genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. Nucleic Acids Res. 2001; 29:4319–4333. [PubMed: 11691919]

Beisel C, Imhof A, Greene J, Kremmer E, Sauer F. Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1. Nature. 2002; 419:857–862. [PubMed: 12397363]

Bernstein BE, Humphrey EL, Erlich RL, Schneider R, Bouman P, Liu JS, Kouzarides T, Schreiber SL. Methylation of histone H3 Lys 4 in coding regions of active genes. Proc Natl Acad Sci U S A. 2002; 99:8695–8700. [PubMed: 12060701]

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell. 2006; 125:315–326. [PubMed: 16630819]

Breen TR. Mutant alleles of the Drosophila trithorax gene produce common and unusual homeotic and other developmental phenotypes. Genetics. 1999; 152:319–344. [PubMed: 10224264]

Byrd KN, Shearn A. ASH1, a Drosophila trithorax group protein, is required for methylation of lysine 4 residues on histone H3. Proc Natl Acad Sci U S A. 2003; 100:11535–11540. [PubMed: 13679578]

Byvoet P, Shepherd GR, Hardin JM, Noland BJ. The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells. Arch Biochem Biophys. 1972; 148:558–567. [PubMed: 5063076]

Capotosti F, Hsieh JJ, Herr W. Species selectivity of mixed-lineage leukemia/trithorax and HCF proteolytic maturation pathways. Mol Cell Biol. 2007; 27:7063–7072. [PubMed: 17698583]

Chinwalla V, Jane EP, Harte PJ. The Drosophila trithorax protein binds to specific chromosomal sites and is co-localized with Polycomb at many sites. Embo J. 1995; 14:2056–2065. [PubMed: 7744011]

Cho YW, Hong T, Hong S, Guo H, Yu H, Kim D, Guszczynski T, Dressler GR, Copeland TD, Kalkum M, Ge K. PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. J Biol Chem. 2007

Chuikov S, Kurash JK, Wilson JR, Xiao B, Justin N, Ivanov GS, McKinney K, Tempst P, Prives C, Gamblin SJ, Barlev NA, Reinberg D. Regulation of p53 activity through lysine methylation. Nature. 2004; 432:353–360. [PubMed: 15525938]

Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell. 2002; 111:185–196. [PubMed: 12408863]

Djabali M, Selleri L, Parry P, Bower M, Young BD, Evans GA. A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukaemias. Nat Genet. 1992; 2:113–118. [PubMed: 1303259]

Dorshkind K, Witte O. Got MLL? Definitive hematopoiesis requires MLL gene expression. Mol Cell. 2004; 13:765–766. [PubMed: 15053868]

Dover J, Schneider J, Tawiah-Boateng MA, Wood A, Dean K, Johnston M, Shilatifard A. Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J Biol Chem. 2002; 277:28368–28371. [PubMed: 12070136]

Duerre JA, Lee CT. In vivo methylation and turnover of rat brain histones. J Neurochem. 1974; 23:541–547. [PubMed: 4421616]

Eissenberg JC, Lee MG, Schneider J, Ilvarsonn A, Shiekhattar R, Shilatifard A. The trithorax-group gene in Drosophila little imaginal discs encodes a trimethylated histone H3 Lys4 demethylase. Nat Struct Mol Biol. 2007a; 14:344–346. [PubMed: 17351630]

Eissenberg JC, Shilatifard A, Dorokhov N, Michener DE. Cdk9 is an essential kinase in Drosophila that is required for heat shock gene expression, histone methylation and elongation factor recruitment. Mol Genet Genomics. 2007b; 277:101–114. [PubMed: 17001490]

Emerling BM, Bonifas J, Kratz CP, Donovan S, Taylor BR, Green ED, Le Beau MM, Shannon KM. MLL5, a homolog of Drosophila trithorax located within a segment of chromosome band 7q22 implicated in myeloid leukemia. Oncogene. 2002; 21:4849–4854. [PubMed: 12101424]

Ernst P, Mabon M, Davidson AJ, Zon LI, Korsmeyer SJ. An Mll-dependent Hox program drives hematopoietic progenitor expansion. Curr Biol. 2004; 14:2063–2069. [PubMed: 15556871]

Fischle W, Franz H, Jacobs SA, Allis CD, Khorasanizadeh S. Specificity of the chromodomain Y chromosome family of chromodomains for lysine-methylated ARK(S/T) motifs. J Biol Chem. 2008; 283:19626–19635. [PubMed: 18450745]

FitzGerald KT, Diaz MO. MLL2: A new mammalian member of the trx/MLL family of genes. Genomics. 1999; 59:187–192. [PubMed: 10409430]

Flanagan JF, Mi LZ, Chruszcz M, Cymborowski M, Clines KL, Kim Y, Minor W, Rastinejad F, Khorasanizadeh S. Double chromodomains cooperate to recognize the methylated histone H3 tail. Nature. 2005; 438:1181–1185. [PubMed: 16372014]

Fujiki R, Chikanishi T, Hashiba W, Ito H, Takada I, Roeder RG, Kitagawa H, Kato S. GlcNAcylation of a histone methyltransferase in retinoic-acid-induced granulopoiesis. Nature. 2009; 459:455–459. [PubMed: 19377461]

Gildea JJ, Lopez R, Shearn A. A screen for new trithorax group genes identified little imaginal discs, the Drosophila melanogaster homologue of human retinoblastoma binding protein 2. Genetics. 2000; 156:645–663. [PubMed: 11014813]

Glaser S, Schaft J, Lubitz S, Vintersten K, van der Hoeven F, Tufteland KR, Aasland R, Anastassiadis K, Ang SL, Stewart AF. Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. Development. 2006; 133:1423–1432. [PubMed: 16540515]

Goo YH, Sohn YC, Kim DH, Kim SW, Kang MJ, Jung DJ, Kwak E, Barlev NA, Berger SL, Chow VT, Roeder RG, Azorsa DO, Meltzer PS, Suh PG, Song EJ, Lee KJ, Lee YC, Lee JW. Activating signal cointegrator 2 belongs to a novel steady-state complex that contains a subset of trithorax group proteins. Mol Cell Biol. 2003; 23:140–149. [PubMed: 12482968]

Gregory GD, Vakoc CR, Rozovskaia T, Zheng X, Patel S, Nakamura T, Canaani E, Blobel GA. Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. Mol Cell Biol. 2007; 27:8466–8479. [PubMed: 17923682]

Gu Y, Nakamura T, Alder H, Prasad R, Canaani O, Cimino G, Croce CM, Canaani E. The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to Drosophila trithorax, to the AF-4 gene. Cell. 1992; 71:701–708. [PubMed: 1423625]

Guenther MG, Jenner RG, Chevalier B, Nakamura T, Croce CM, Canaani E, Young RA. Global and Hox-specific roles for the MLL1 methyltransferase. Proc Natl Acad Sci U S A. 2005; 102:8603–8608. [PubMed: 15941828]

Heuser M, Yap DB, Leung M, de Algara TR, Tafech A, McKinney S, Dixon J, Thresher R, Colledge B, Carlton M, Humphries RK, Aparicio SA. Loss of MLL5 results in pleiotropic hematopoietic defects, reduced neutrophil immune function, and extreme sensitivity to DNA demethylation. Blood. 2009; 113:1432–1443. [PubMed: 18854576]

Honda BM, Candido PM, Dixon GH. Histone methylation. Its occurrence in different cell types and relation to histone H4 metabolism in developing trout testis. J Biol Chem. 1975; 250:8686–8689. [PubMed: 1184586]

Honma T, Goto K. Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature. 2001; 409:525–529. [PubMed: 11206550]

Hsieh JJ, Ernst P, Erdjument-Bromage H, Tempst P, Korsmeyer SJ. Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. Mol Cell Biol. 2003; 23:186–194. [PubMed: 12482972]

Hughes CM, Rozenblatt-Rosen O, Milne TA, Copeland TD, Levine SS, Lee JC, Hayes DN, Shanmugam KS, Bhattacharjee A, Biondi CA, Kay GF, Hayward NK, Hess JL, Meyerson M. Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. Mol Cell. 2004; 13:587–597. [PubMed: 14992727]

Huntsman DG, Chin SF, Muleris M, Batley SJ, Collins VP, Wiedemann LM, Aparicio S, Caldas C. MLL2, the second human homolog of the Drosophila trithorax gene, maps to 19q13.1 and is amplified in solid tumor cell lines. Oncogene. 1999; 18:7975–7984. [PubMed: 10637508]

Ingham PW. A clonal analysis of the requirement for the trithorax gene in the diversification of segments in Drosophila. J Embryol Exp Morphol. 1985; 89:349–365. [PubMed: 4093752]

Ingham PW. trithorax and the regulation of homeotic gene expression in Drosophila: a historical perspective. Int J Dev Biol. 1998; 42:423–429. [PubMed: 9654027]

Ingham PW, Whittle R. Trithorax: a new homeotic mutation of Drosophila melanogaster causing transformations of abdominal and thoracic imaginal segments. Molecular and General Genetics. 1980; 179:607–614.

Katsani KR, Arredondo JJ, Kal AJ, Verrijzer CP. A homeotic mutation in the trithorax SET domain impedes histone binding. Genes Dev. 2001; 15:2197–2202. [PubMed: 11544176]

Kelley DE, Stokes DG, Perry RP. CHD1 interacts with SSRP1 and depends on both its chromodomain and its ATPase/helicase-like domain for proper association with chromatin. Chromosoma. 1999; 108:10–25. [PubMed: 10199952]

Klymenko T, Muller J. The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. EMBO Rep. 2004; 5:373–377. [PubMed: 15031712]

Konev AY, Tribus M, Park SY, Podhraski V, Lim CY, Emelyanov AV, Vershilova E, Pirrotta V, Kadonaga JT, Lusser A, Fyodorov DV. CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin in vivo. Science. 2007; 317:1087–1090. [PubMed: 17717186]

Kornberg RD. Chromatin structure: a repeating unit of histones and DNA. Science. 1974; 184:868–871. [PubMed: 4825889]

Kornberg RD, Lorch Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell. 1999; 98:285–294. [PubMed: 10458604]

Kouskouti A, Scheer E, Staub A, Tora L, Talianidis I. Gene-specific modulation of TAF10 function by SET9-mediated methylation. Mol Cell. 2004; 14:175–182. [PubMed: 15099517]

Kouzarides T. Chromatin modifications and their function. Cell. 2007; 128:693–705. [PubMed: 17320507]

Kratz CP, Emerling BM, Donovan S, Laig-Webster M, Taylor BR, Thompson P, Jensen S, Banerjee A, Bonifas J, Makalowski W, Green ED, Le Beau MM, Shannon KM. Candidate gene isolation and comparative analysis of a commonly deleted segment of 7q22 implicated in myeloid malignancies. Genomics. 2001; 77:171–180. [PubMed: 11597142]

Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. Nat Rev Cancer. 2007; 7:823–833. [PubMed: 17957188]

Krogan NJ, Dover J, Khorrami S, Greenblatt JF, Schneider J, Johnston M, Shilatifard A. COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression. J Biol Chem. 2002; 277:10753–10755. [PubMed: 11805083]

Krogan NJ, Dover J, Wood A, Schneider J, Heidt J, Boateng MA, Dean K, Ryan OW, Golshani A, Johnston M, Greenblatt JF, Shilatifard A. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol Cell. 2003; 11:721–729. [PubMed: 12667454]

Le Beau MM, Espinosa R 3rd, Davis EM, Eisenbart JD, Larson RA, Green ED. Cytogenetic and molecular delineation of a region of chromosome 7 commonly deleted in malignant myeloid diseases. Blood. 1996; 88:1930–1935. [PubMed: 8822909]

Lee J, Saha PK, Yang QH, Lee S, Park JY, Suh Y, Lee SK, Chan L, Roeder RG, Lee JW. Targeted inactivation of MLL3 histone H3-Lys-4 methyltransferase activity in the mouse reveals vital roles for MLL3 in adipogenesis. Proc Natl Acad Sci U S A. 2008; 105:19229–19234. [PubMed: 19047629]

Lee JH, Skalnik DG. CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex. J Biol Chem. 2005; 280:41725–41731. [PubMed: 16253997]

Lee JH, Tate CM, You JS, Skalnik DG. Identification and characterization of the human Set1B histone H3-Lys4 methyltransferase complex. J Biol Chem. 2007a; 282:13419–13428. [PubMed: 17355966]

Lee N, Zhang J, Klose RJ, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. The trithorax-group protein Lid is a histone H3 trimethyl-Lys4 demethylase. Nat Struct Mol Biol. 2007b; 14:341–343. [PubMed: 17351631]

Lee S, Lee DK, Dou Y, Lee J, Lee B, Kwak E, Kong YY, Lee SK, Roeder RG, Lee JW. Coactivator as a target gene specificity determinant for histone H3 lysine 4 methyltransferases. Proc Natl Acad Sci U S A. 2006; 103:15392–15397. [PubMed: 17021013]

Lu X, Simon MD, Chodaparambil JV, Hansen JC, Shokat KM, Luger K. The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. Nat Struct Mol Biol. 2008; 15:1122–1124. [PubMed: 18794842]

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997; 389:251–260. [PubMed: 9305837]

Luna-Fineman S, Shannon KM, Lange BJ. Childhood monosomy 7: epidemiology, biology, and mechanistic implications. Blood. 1995; 85:1985–1999. [PubMed: 7718870]

Madan V, Madan B, Brykczynska U, Zilbermann F, Hogeveen K, Dohner K, Dohner H, Weber O, Blum C, Rodewald HR, Sassone-Corsi P, Peters AH, Fehling HJ. Impaired function of primitive hematopoietic cells in mice lacking the Mixed-Lineage-Leukemia homolog MLL5. Blood. 2009; 113:1444–1454. [PubMed: 18952892]

Mazo AM, Huang DH, Mozer BA, Dawid IB. The trithorax gene, a trans-acting regulator of the bithorax complex in Drosophila, encodes a protein with zinc-binding domains. Proc Natl Acad Sci U S A. 1990; 87:2112–2116. [PubMed: 2107543]

Mbangkollo D, Burnett R, McCabe N, Thirman M, Gill H, Yu H, Rowley JD, Diaz MO. The human MLL gene: nucleotide sequence, homology to the Drosophila trx zinc-finger domain, and alternative splicing. DNA Cell Biol. 1995; 14:475–483. [PubMed: 7598802]

McMahon KA, Hiew SY, Hadjur S, Veiga-Fernandes H, Menzel U, Price AJ, Kioussis D, Williams O, Brady HJ. Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. Cell Stem Cell. 2007; 1:338–345. [PubMed: 18371367]

Miller T, Krogan NJ, Dover J, Erdjument-Bromage H, Tempst P, Johnston M, Greenblatt JF, Shilatifard A. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. Proc Natl Acad Sci U S A. 2001; 98:12902–12907. [PubMed: 11687631]

Mo R, Rao SM, Zhu YJ. Identification of the MLL2 complex as a coactivator for estrogen receptor alpha. J Biol Chem. 2006; 281:15714–15720. [PubMed: 16603732]

Murray K. The Occurrence of Epsilon-N-Methyl Lysine in Histones. Biochemistry. 1964; 3:10–15. [PubMed: 14114491]

Nakamura T, Blechman J, Tada S, Rozovskaia T, Itoyama T, Bullrich F, Mazo A, Croce CM, Geiger B, Canaani E. huASH1 protein, a putative transcription factor encoded by a human homologue of the Drosophila ash1 gene, localizes to both nuclei and cell-cell tight junctions. Proc Natl Acad Sci U S A. 2000; 97:7284–7289. [PubMed: 10860993]

Ng DW, Wang T, Chandrasekharan MB, Aramayo R, Kertbundit S, Hall TC. Plant SET domain-containing proteins: structure, function and regulation. Biochim Biophys Acta. 2007; 1769:316–329. [PubMed: 17512990]

Ng HH, Robert F, Young RA, Struhl K. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol Cell. 2003; 11:709–719. [PubMed: 12667453]

Nislow C, Ray E, Pillus L. SET1, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes. Mol Biol Cell. 1997; 8:2421–2436. [PubMed: 9398665]

Ono R, Nosaka T, Hayashi Y. Roles of a trithorax group gene, MLL, in hematopoiesis. Int J Hematol. 2005; 81:288–293. [PubMed: 15914356]

Pavri R, Zhu B, Li G, Trojer P, Mandal S, Shilatifard A, Reinberg D. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. Cell. 2006; 125:703–717. [PubMed: 16713563]

Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF. B and C floral organ identity functions require SEPALLATA MADS-box genes. Nature. 2000; 405:200–203. [PubMed: 10821278]

Prasad R, Zhadanov AB, Sedkov Y, Bullrich F, Druck T, Rallapalli R, Yano T, Alder H, Croce CM, Huebner K, Mazo A, Canaani E. Structure and expression pattern of human ALR, a novel gene with strong homology to ALL-1 involved in acute leukemia and to Drosophila trithorax. Oncogene. 1997; 15:549–560. [PubMed: 9247308]

Riechmann JL, Meyerowitz EM. MADS domain proteins in plant development. Biol Chem. 1997; 378:1079–1101. [PubMed: 9372178]

Roguev A, Schaft D, Shevchenko A, Pijnappel WW, Wilm M, Aasland R, Stewart AF. The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. Embo J. 2001; 20:7137–7148. [PubMed: 11742990]

Rowley JD. The critical role of chromosome translocations in human leukemias. Annu Rev Genet. 1998; 32:495–519. [PubMed: 9928489]

Rozovskaia T, Tillib S, Smith S, Sedkov Y, Rozenblatt-Rosen O, Petruk S, Yano T, Nakamura T, Ben-Simchon L, Gildea J, Croce CM, Shearn A, Canaani E, Mazo A. Trithorax and ASH1 interact directly and associate with the trithorax group-responsive bxd region of the Ultrabithorax promoter. Mol Cell Biol. 1999; 19:6441–6447. [PubMed: 10454589]

Ruault M, Brun ME, Ventura M, Roizes G, De Sario A. MLL3, a new human member of the TRX/MLL gene family, maps to 7q36, a chromosome region frequently deleted in myeloid leukaemia. Gene. 2002; 284:73–81. [PubMed: 11891048]

Rudolph T, Yonezawa M, Lein S, Heidrich K, Kubicek S, Schafer C, Phalke S, Walther M, Schmidt A, Jenuwein T, Reuter G. Heterochromatin formation in Drosophila is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3-3. Mol Cell. 2007; 26:103–115. [PubMed: 17434130]

Saleh A, Al-Abdallat A, Ndamukong I, Alvarez-Venegas R, Avramova Z. The Arabidopsis homologs of trithorax (ATX1) and enhancer of zeste (CLF) establish 'bivalent chromatin marks' at the silent AGAMOUS locus. Nucleic Acids Res. 2007; 35:6290–6296. [PubMed: 17881378]

Sanchez-Elsner T, Gou D, Kremmer E, Sauer F. Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. Science. 2006; 311:1118–1123. [PubMed: 16497925]

Schneider J, Dover J, Johnston M, Shilatifard A. Global proteomic analysis of S. cerevisiae (GPS) to identify proteins required for histone modifications. Methods Enzymol. 2004; 377:227–234. [PubMed: 14979028]

Schuettengruber B, Ganapathi M, Leblanc B, Portoso M, Jaschek R, Tolhuis B, van Lohuizen M, Tanay A, Cavalli G. Functional anatomy of polycomb and trithorax chromatin landscapes in Drosophila embryos. PLoS Biol. 2009; 7:e13. [PubMed: 19143474]

Secombe J, Li L, Carlos L, Eisenman RN. The Trithorax group protein Lid is a trimethyl histone H3K4 demethylase required for dMyc-induced cell growth. Genes Dev. 2007; 21:537–551. [PubMed: 17311883]

Sedkov Y, Benes JJ, Berger JR, Riker KM, Tillib S, Jones RS, Mazo A. Molecular genetic analysis of the Drosophila trithorax-related gene which encodes a novel SET domain protein. Mech Dev. 1999; 82:171–179. [PubMed: 10354481]

Sedkov Y, Cho E, Petruk S, Cherbas L, Smith ST, Jones RS, Cherbas P, Canaani E, Jaynes JB, Mazo A. Methylation at lysine 4 of histone H3 in ecdysone-dependent development of Drosophila. Nature. 2003; 426:78–83. [PubMed: 14603321]

Sedkov Y, Tillib S, Mizrokhi L, Mazo A. The bithorax complex is regulated by trithorax earlier during Drosophila embryogenesis than is the Antennapedia complex, correlating with a bithorax-like expression pattern of distinct early trithorax transcripts. Development. 1994; 120:1907–1917. [PubMed: 7924996]

Shearn A. What is the normal function of genes which give rise to homeotic mutations? Basic Life Sci. 1980; 16:155–162. [PubMed: 7458795]

Shearn A. The ash-1, ash-2 and trithorax genes of Drosophila melanogaster are functionally related. Genetics. 1989; 121:517–525. [PubMed: 2497049]

Shearn A, Hersperger E, Hersperger G. Genetic studies of mutations at two loci of Drosophila melanogaster which cause a wide variety of homeotic transformations. Roux's archives of developmental biology. 1987; 196:231–242.

Shearn A, Rice T, Garen A, Gehring W. Imaginal disc abnormalities in lethal mutants of Drosophila. Proc Natl Acad Sci U S A. 1971; 68:2594–2598. [PubMed: 5002822]

Shi Y, Whetstine JR. Dynamic regulation of histone lysine methylation by demethylases. Mol Cell. 2007; 25:1–14. [PubMed: 17218267]

Shilatifard A. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu Rev Biochem. 2006; 75:243–269. [PubMed: 16756492]

Shilatifard A. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. Curr Opin Cell Biol. 2008; 20:341–348. [PubMed: 18508253]

Sims RJ 3rd, Chen CF, Santos-Rosa H, Kouzarides T, Patel SS, Reinberg D. Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. J Biol Chem. 2005; 280:41789–41792. [PubMed: 16263726]

Smith E, Shilatifard A. Developmental biology. Histone cross-talk in stem cells. Science. 2009; 323:221–222. [PubMed: 19131622]

Smith ER, Lee MG, Winter B, Droz NM, Eissenberg JC, Shiekhattar R, Shilatifard A. Drosophila UTX is a histone H3 Lys27 demethylase that colocalizes with the elongating form of RNA polymerase II. Mol Cell Biol. 2008; 28:1041–1046. [PubMed: 18039863]

Smith ST, Petruk S, Sedkov Y, Cho E, Tillib S, Canaani E, Mazo A. Modulation of heat shock gene expression by the TAC1 chromatin-modifying complex. Nat Cell Biol. 2004; 6:162–167. [PubMed: 14730313]

Stassen MJ, Bailey D, Nelson S, Chinwalla V, Harte PJ. The Drosophila trithorax proteins contain a novel variant of the nuclear receptor type DNA binding domain and an ancient conserved motif found in other chromosomal proteins. Mech Dev. 1995; 52:209–223. [PubMed: 8541210]

Stokes DG, Tartof KD, Perry RP. CHD1 is concentrated in interbands and puffed regions of Drosophila polytene chromosomes. Proc Natl Acad Sci U S A. 1996; 93:7137–7142. [PubMed: 8692958]

Tan YC, Chow VT. Novel human HALR (MLL3) gene encodes a protein homologous to ALR and to ALL-1 involved in leukemia, and maps to chromosome 7q36 associated with leukemia and developmental defects. Cancer Detect Prev. 2001; 25:454–469. [PubMed: 11718452]

Tenney K, Gerber M, Ilvarsonn A, Schneider J, Gause M, Dorsett D, Eissenberg JC, Shilatifard A. Drosophila Rtf1 functions in histone methylation, gene expression, and Notch signaling. Proc Natl Acad Sci U S A. 2006; 103:11970–11974. [PubMed: 16882721]

Tenney K, Shilatifard A. A COMPASS in the voyage of defining the role of trithorax/MLL-containing complexes: linking leukemogensis to covalent modifications of chromatin. J Cell Biochem. 2005; 95:429–436. [PubMed: 15786493]

Terranova R, Agherbi H, Boned A, Meresse S, Djabali M. Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll. Proc Natl Acad Sci U S A. 2006; 103:6629–6634. [PubMed: 16618927]

Tkachuk DC, Kohler S, Cleary ML. Involvement of a homolog of Drosophila trithorax by 11q23 chromosomal translocations in acute leukemias. Cell. 1992; 71:691–700. [PubMed: 1423624]

Tripoulas N, LaJeunesse D, Gildea J, Shearn A. The Drosophila ash1 gene product, which is localized at specific sites on polytene chromosomes, contains a SET domain and a PHD finger. Genetics. 1996; 143:913–928. [PubMed: 8725238]

Trojer P, Reinberg D. Histone lysine demethylases and their impact on epigenetics. Cell. 2006; 125:213–217. [PubMed: 16630806]

Wang PF, Lin C, Simth ER, Sanderson BW, Wu M, Gogol M, Alexander T, Seidel C, Wiedemann LM, Guo H, Krumlauf R, Shilatifard A. Global Analysis of H3K4 methylation defines Mll1 (KMT2A) as a gene-specific activator of transcription. Submitted.

Waterborg JH. Dynamic methylation of alfalfa histone H3. J Biol Chem. 1993; 268:4918–4921. [PubMed: 8444870]

Wood A, Krogan NJ, Dover J, Schneider J, Heidt J, Boateng MA, Dean K, Golshani A, Zhang Y, Greenblatt JF, Johnston M, Shilatifard A. Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. Mol Cell. 2003a; 11:267–274. [PubMed: 12535539]

Wood A, Schneider J, Dover J, Johnston M, Shilatifard A. 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. 2003b; 278:34739–34742. [PubMed: 12876294]

Wood A, Schneider J, Dover J, Johnston M, Shilatifard A. The Bur1/Bur2 complex is required for histone H2B monoubiquitination by Rad6/Bre1 and histone methylation by COMPASS. Mol Cell. 2005; 20:589–599. [PubMed: 16307922]

Wood A, Shukla A, Schneider J, Lee JS, Stanton JD, Dzuiba T, Swanson SK, Florens L, Washburn MP, Wyrick J, Bhaumik SR, Shilatifard A. Ctk complex-mediated regulation of histone methylation by COMPASS. Mol Cell Biol. 2007; 27:709–720. [PubMed: 17088385]

Wu M, Wang PF, Lee JS, Martin-Brown S, Florens L, Washburn M, Shilatifard A. Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. Mol Cell Biol. 2008; 28:7337–7344. [PubMed: 18838538]

Yu BD, Hanson RD, Hess JL, Horning SE, Korsmeyer SJ. MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. Proc Natl Acad Sci U S A. 1998; 95:10632–10636. [PubMed: 9724755]

Yu BD, Hess JL, Horning SE, Brown GA, Korsmeyer SJ. Altered Hox expression and segmental identity in Mll-mutant mice. Nature. 1995; 378:505–508. [PubMed: 7477409]

Zhang Y, Wong J, Klinger M, Tran MT, Shannon KM, Killeen N. MLL5 contributes to hematopoietic stem cell fitness and homeostasis. Blood. 2009; 113:1455–1463. [PubMed: 18818388]

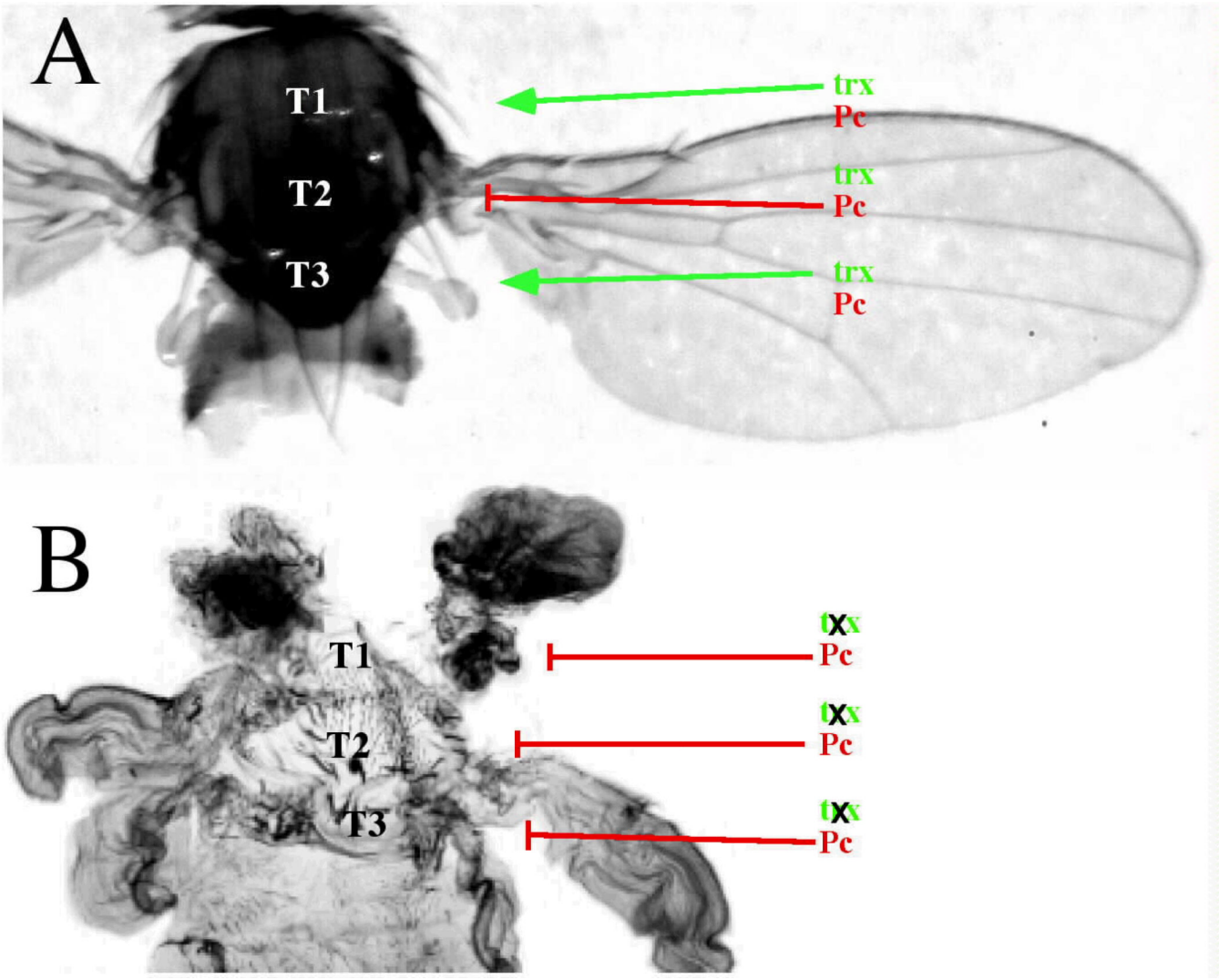

Figure 1. *Trithorax* and the specification of axial segment identity

A. The three segments of the wild-type fly thorax are decorated with distinct structures; from anterior to posterior, the prothoracic segment (T1) carries a ventral pair of legs, the mesothoracic segment (T2) carries both a ventral pair of legs and a dorsal pair of wings, and the metathoracic segment (T3) carries a ventral pair of legs and a dorsal pair of vestigial wings called halteres. In the prothoracic and metathoracic segments, *trithorax* ("trx") prevents the silencing of genes that are repressed in the mesothroax by the *Polycomb* group ("Pc") repressors. **B**. In flies lacking *trithorax* function, *Polycomb* group repressors are unopposed in the pro- and metathoracic segments, causing these segments to develop as mesothoracic segments, with a dorsal pair of wings on each segment. Panel B adapted from Ingham, P.W. (1998) Int. J. Dev. Biol. 42, 423–429. Used by permission.

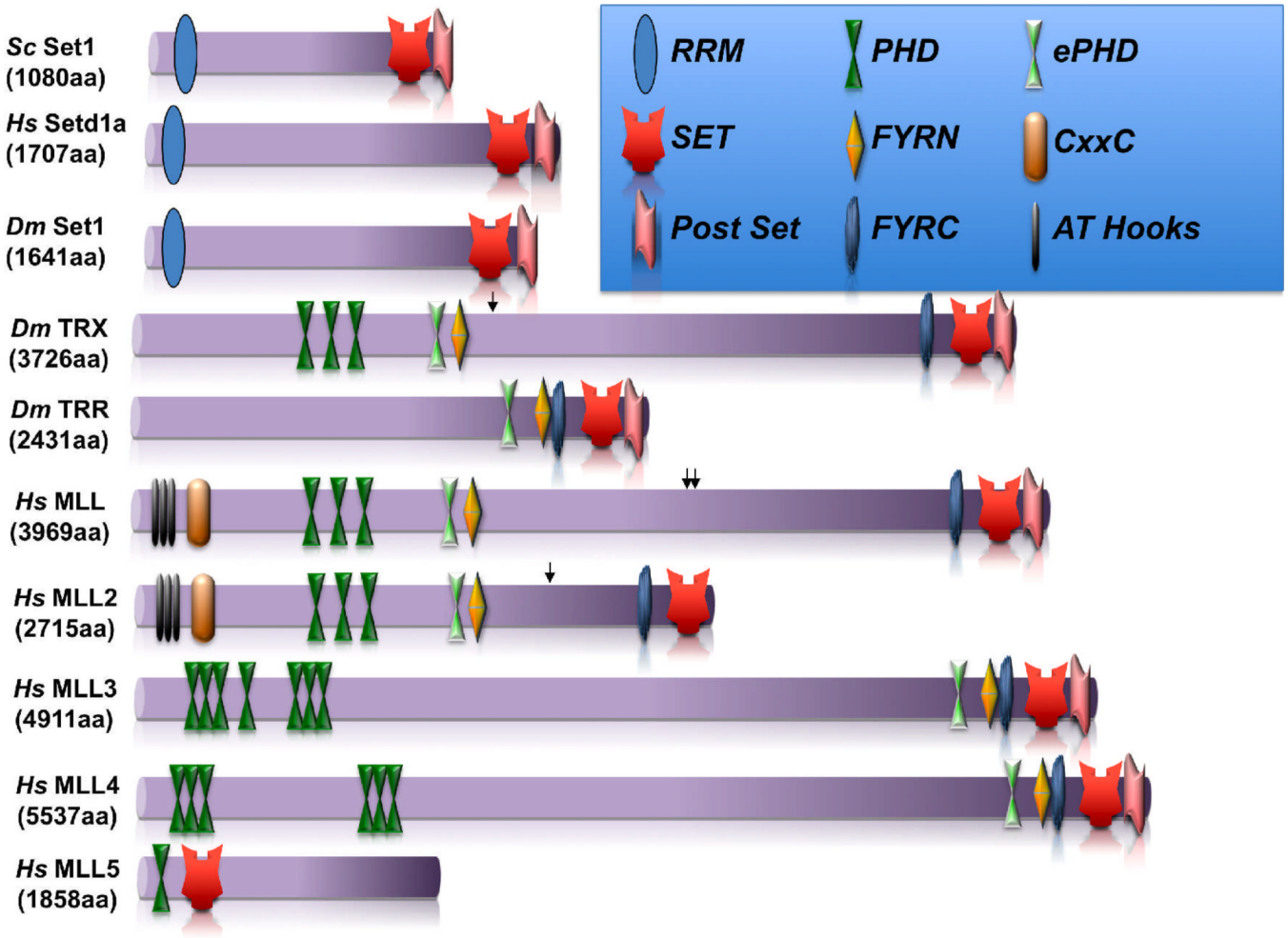

Figure 2. Domain structure of animal H3K4 methylases

The domain organization for the founding member of the H3K4 methylase, the yeast Set1 protein, and its homolog and related proteins in Drosophila (Set1, TRX and TRR) and human (Set1, and MLL1–5) are shown above. Numbers to the left of each structure diagram indicate protein length in amino acids. Arrows indicate the proteolytic cleavage sites on each protein.

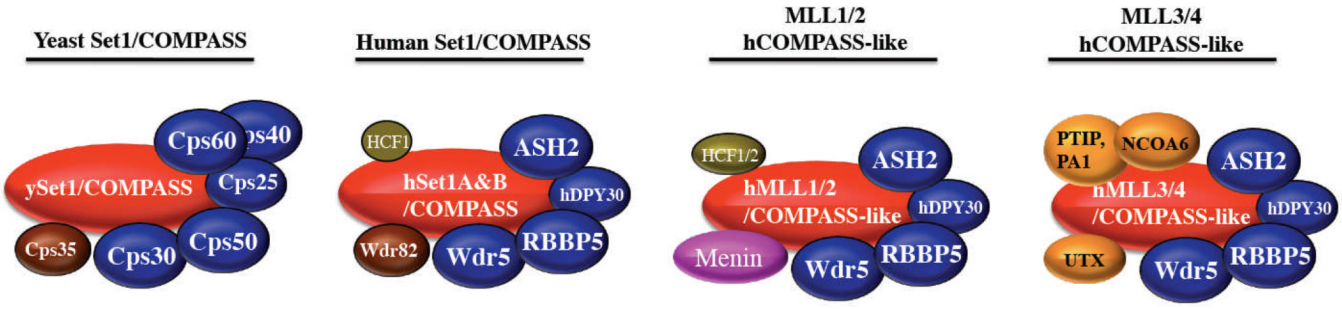

Figure 3. Subunit composition for yeast COMPASS and its mammalian homologues Set1 A/B and MLL1–4 complexes

The founding member of the H3K4 methyltransferases, Set1/COMPASS was identified in yeast. Human hSet1 (hSet1A and B) and MLLs (MLL1–4) are also found in COMPASS and COMPASS-like complexes, respectively. Each complex is capable of methylating histone H3 on its fourth lysine (H3K4). The known common subunits shared between yeast and mammalian complexes are shown in BLUE B. Cps35 in yeast and its homologue in human Wdr82 are found only in COMPASS and Set1 A/B complexes. This subunit is shown in BROWN. Menin, which is subunit shared only in the MLL1/2 complexes is shown in LAVENDER. MLL3/4 specific components are shown in GOLD.

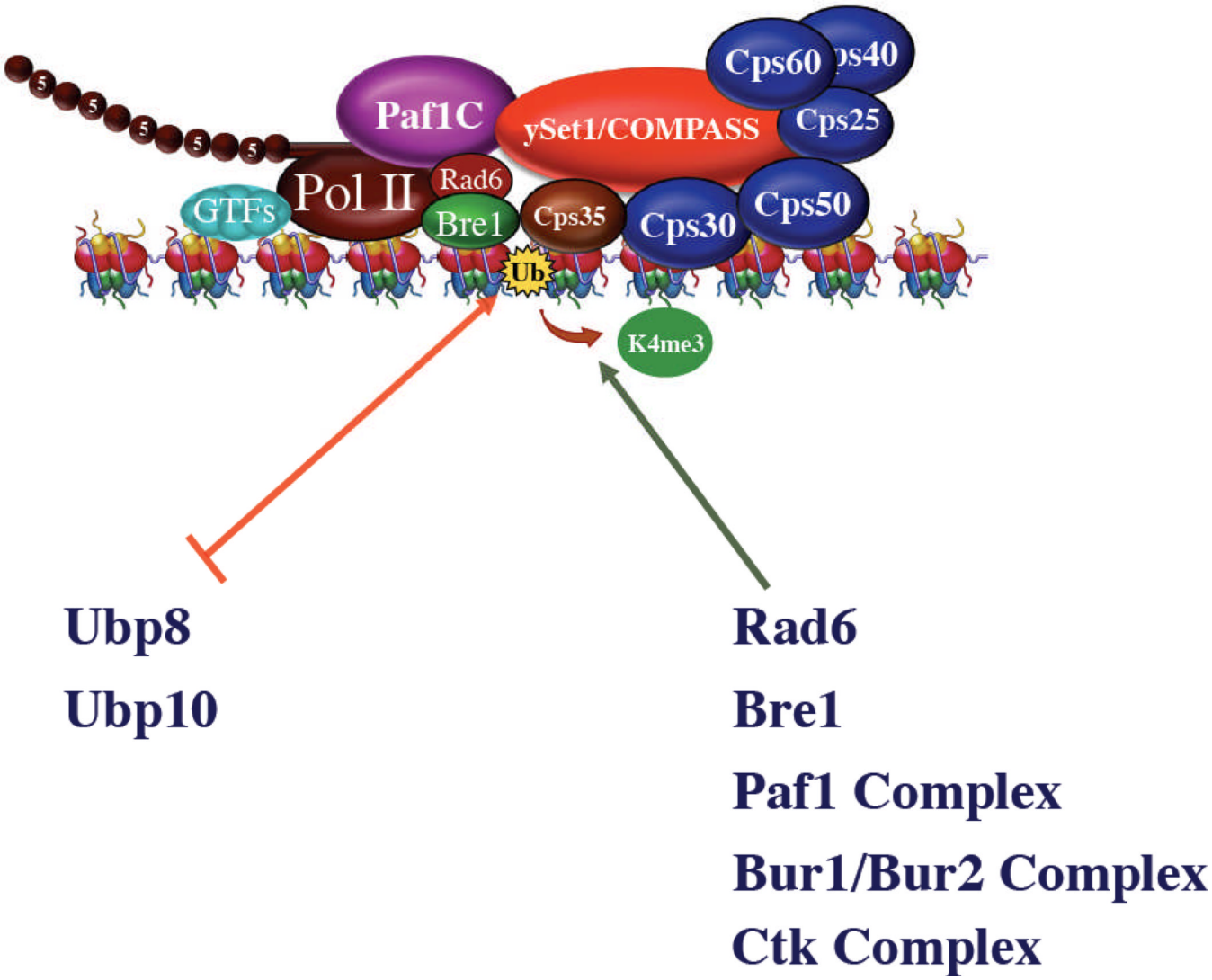

Figure 4. Schematic representation of the molecular machinery uncovered via a biochemical screen, GPS, that are required for proper H3K4 methylation by COMPASS

Based on our work, as well as results from other laboratories, we know that COMPASS is required for methylation of H3K4. GPS screen uncovered a role for the Rad6/Bre1 complex in monoubiquitination of histone H2B, which signals the methylation of H3K4 by COMPASS. The Cps35 subunit of COMPASS interacts with chromatin in an H2B monoubiquitination-dependnet manner resulting in the assembly of trimethylation competent COMPASS. Other factors such as the elongation factor, the Paf1 complex, and protein kinases Bur1/Bur2 complex and Ctk complex were identified through our GPS screen as factor required for proper H3K4 methylation. Much of the machinery identified in yeast functioning in H2B monoubiquitination or H2B deubiquitination and H3K4 methylation are highly conserved, both structurally and functionally, from yeast to human.