Biochemistry and Cell Biology
Biochimie et biologie cellulaire

Histone H3K4 trimethylation: dynamic interplay with pre-mRNA splicing

Journal:	*Biochemistry and Cell Biology*
Manuscript ID:	bcb-2015-0065.R1
Manuscript Type:	Mini Review
Date Submitted by the Author:	21-Jun-2015
Complete List of Authors:	Davie, James; University of Manitoba, Xu, Wayne; University of Manitoba, Biochemistry and Medical Genetics Delcuve, Genevieve; University of Manitoba, Biochemistry and Medical Genetics
Keyword:	histone H3 lysine 4 trimethylation, gene expression, epigenetics, pre-mRNA splicing

Histone H3K4 trimethylation: dynamic interplay with pre-mRNA splicing

James R Davie, Wayne Xu and Genevieve P Delcuve

Children's Hospital Research Institute of Manitoba, University of Manitoba,
Winnipeg, Manitoba, Canada R3E 3P4

Address for Correspondence:

Dr. James R. Davie
Children's Hospital Research Institute of Manitoba
715 McDermot Avenue
Winnipeg, Manitoba, R3E 3P4
Tel. # (204)-975-7732
Fax # (204)-977-5691
E-mail: jim.davie@umanitoba.ca

Draft

Abstract

Histone H3 lysine 4 trimethylation (H3K4me3) is often stated as a mark of transcriptionally active promoters. However, closer study of the positioning of H3K4me3 shows the mark locating primarily after the first exon at the 5' splice site and overlapping with a CpG island in mammalian cells. There are several enzyme complexes that are involved in the placement of the H3K4me3 mark including multiple protein complexes containing SETD1A, SETD1B, MLL1 and MLL2 enzymes (writers). CXXC1, which is associated with SETD1A and SETD1B, target these enzymes to unmethylated CpG islands. Lysine demethylases (KDM5 family members, erasers) demethylate H3K4me3. The H3K4me3 mark is recognized by several proteins (readers), including lysine acetyltransferase complexes, chromatin remodelers, and RNA bound proteins involved in pre-mRNA splicing. Interestingly, attenuation of H3K4me3 impacts pre-mRNA splicing, and inhibition of pre-mRNA splicing attenuates H3K4me3.

Key words: histone H3 lysine 4 trimethylation, gene expression, epigenetics, pre-mRNA splicing

Draft

INTRODUCTION

In human cells, 95% of the multi-exon genes yield alternatively spliced transcripts, which can give rise to proteins with different functions and cellular locations (Khan et al., 2012, Kornblihtt et al., 2013, Ip et al., 2011, Hnilicova and Stanek, 2011). Pre-mRNA splicing occurs co-transcriptionally in mammalian cells (Sperling et al., 2008, Wahl et al., 2009, Kalsotra and Cooper, 2011). The regulation of alternative splicing is quite complex and is influenced by transcription rates, histone post-translational modifications (**PTMs**) and a variety of splicing factors (Braunschweig et al., 2013, Kornblihtt et al., 2013, Bentley, 2014). The rate of elongation through a gene is not uniform and is influenced by histone PTMs (Kwak et al., 2013, Zhou et al., 2014, Dujardin et al., 2013). The types of histone PTMs located along a transcribed gene body are polarized. H3K4me3 is typically found at the 5' end of the gene body, followed by H3K36me3 further down the gene. In this review, we will focus on the relationship between H3K4me3 location at the beginning of the transcribed gene body and pre-mRNA splicing.

Draft

Histone modifications and chromatin structure/function

Nuclear DNA is packaged into nucleosomes, the basic repeating structural units in chromatin. The nucleosome consists of a histone octamer, arranged as a $(H3\text{-}H4)_2$ tetramer and two H2A-H2B dimers, around which DNA is wrapped. The core histones (H2A, H2B, H3, H4) have a similar structure with a basic N terminal domain, a globular domain organized by the histone fold, and a C terminal tail.

The N terminal tails emanate from the nucleosome in all directions and interact with linker DNA, nearby nucleosomes or with other proteins. Core histones undergo a variety of reversible PTMs, including acetylation, methylation and phosphorylation (Maze et al., 2014). Histone PTMs are key players in the epigenetic regulation of gene expression. Histone PTMs are added to specific sites on the core histones by a variety of enzymes (***"writers"***) (e.g. lysine acetyltransferases, **KATs**; lysine methyltransferases, **KMTs**), and removed by specific enzymes (***"erasers"***) (e.g. histone deacetylases, **HDACs**; lysine demethylases, **KDMs**) (Allis et al., 2007). Some histone PTMs (active marks) are associated with transcribed chromatin regions, while others (repressive marks) are present in silent regions. Histone acetylation, H3K4me3 and H3K36me3 are active gene marks, whereas H3K9me3 and H3K27me3 are repressive marks. Histone PTMs function to alter chromatin structure and/or provide a "code" for recruitment or occlusion of nonhistone chromosomal proteins to chromatin. These recruited proteins are referred to as "***readers***". Thus the nucleosome is more than a structural unit; it is also a signalling unit (Turner, 2014).

Draft

Pre-mRNA splicing, nucleosomes and transcription

Most human genes have exons separated by introns. The introns are removed from the transcript to join the exons together by a process called splicing (Figure 1) to produce a mature RNA that is translated to synthesize a protein (Sperling et al., 2008). Pre-mRNA splicing is a regulated stepwise process catalyzed by a multicomponent machine (the spliceosome) consisting of

RNA and protein, and this process occurs co-transcriptionally in mammalian cells (Sperling et al., 2008, Wahl et al., 2009, Kalsotra and Cooper, 2011).

There is a functional connection between chromatin organization and pre-mRNA splicing (Luco et al., 2010, Khan et al., 2012, Sims et al., 2007, de Almeida and Carmo-Fonseca, 2014). Nucleosomes are non-uniformly distributed along the body of transcribed genes as they are enriched on exons relative to introns and have preferential positioning at exon-intron and intron-exon boundaries (Chodavarapu et al., 2010, Tilgner et al., 2009). Further, most internal exons are short, about the length of nucleosomal DNA (Tilgner et al., 2009, Schwartz et al., 2009).

Draft

H3K4me3 genomic location

Typically H3K4me3 is viewed as a mark of an active promoter. However, in mammalian cells the predominant H3K4me3 peak is located at the end of the first exon at the site of the 5' splice site (Bieberstein et al., 2012) (Figure 1). To reach this conclusion, the authors used H3K4me3 chromatin immunoprecipitation (**ChIP**) sequencing data from the human ENCODE project (EncodeProjectConsortium, 2012, EncodeProjectConsortium, 2011). Current evidence shows that the location of the H3K4me3 modified nucleosomes along a transcribed gene body depends on 1) the location of a CpG island (unmethylated state) and 2) pre-mRNA splicing (Chen et al., 2014, Bieberstein et al., 2012) (Figure 1).

H3K4me3 is highly biased towards genes with CpG islands, which are usually located at the 5' end of the gene body (Pal et al., 2011, Chen et al., 2014, Maunakea et al., 2010, Barrera et al., 2008) (Figure 1 and 2). H3K4me3 is not present at methylated CpG islands (Deaton et al., 2011). Artificial CpG clusters introduced into the mouse genome at sites that normally lack H3K4me3 became associated with peaks of H3K4me3 modified nucleosomes in the absence of a promoter or any associated RNA polymerase (Thomson et al., 2010). This result presents evidence that transcription is not a prerequisite for the establishment of H3K4me3. Although H3K4me3 locates preferentially to CpG islands, genes lacking CpG islands in their 5' region may also form peaks of H3K4me3 (Barrera et al., 2008). Figure 3 shows the *APOL1*, *APOL2* and *TFF3* genes, which appear to lack CpG islands, have an H3K4me3 peak located at the end of the first exon in HepG2 cells. Note that for the *APOL2* gene the H3K4me3 peak is located at the end of the first exon in HepG2, A549 and HCT116 cells.

Draft

As shown in Figures 2 and 3, H3K4me3 is often located in several peaks. The major peak(s) is located downstream to the transcription start site (**TSS**), while a lesser peak is upstream to the TSS aligning with the upstream promoter region of the gene (Barrera et al., 2008, Barski et al., 2007, Guenther et al., 2007, Rahl et al., 2010, Bieberstein et al., 2012). Figure 2 shows that the *RBM23* and *PRMT5* genes have a minor peak of H3K4me3 upstream of the TSS and a major peak of H3K4me3 downstream of the TSS at the end of the first exon in HCT116 and A549 cells. The immediate-early gene *FOS* has one predominant peak of H3K4me3 at the end of exon 1, while the immediate-early gene *DUSP1* has two

peaks of H3K4me3 with one peak positioned at the end of exon 1. The human *MCL1* has one peak of H3K4me3 at the exon 1/intron 1 boundary that overlaps with a CpG island, while another peak of H3K4me3 is located over exon 2 which is downstream of the CpG island (Figure 1 and 2). The locations of the H3K4me3 peaks downstream of the *MCL1* TSS are similar in HCT116, HeLa, A549 and MCF7 cells (EncodeProjectConsortium, 2011). The two peaks of H3K4me3 cover 2 to 3 nucleosomes. Currently the processes involved in establishing multiple peaks of H3K4me3 in the gene body are not known.

However, there are exceptions, that is, H3K4me3 is not always positioned at the end of exon 1 (Deaton et al., 2011). The human *MPL* gene codes for a type 1 hematopoietic cytokine receptor as well a truncated protein when the transcript is alternatively spliced at exons 9 and 10 (Xiao et al., 2014). The H3K4me3 modified nucleosomes are located well down the gene body at alternative exons 9 and 10 in HCT116, A549, HeLa and MCF7 cells (Figure 4). The H3K4me3 peaks are found downstream of H3K36me3, which is typically found in the body of transcribed genes. Of note, there is a CpG island positioned at the H3K4me3 peaks, consistent with the idea that the positioning of the H3K4me3 mark is guided by the CpG island.

H3K4me3 modified nucleosomes are often located at the first 5' splice site of transcribed mammalian genes, with the location of the H3K4me3 being dependent on pre-mRNA splicing (Bieberstein et al., 2012). Inhibition of pre-mRNA splicing resulted in the loss of the H3K4me3 located at the exon 1 5' splice site (Bieberstein et al., 2012). Collectively, these reports suggest that the

Draft

CpG island location and splicing play roles in deciding which nucleosome(s) will be trimethylated at H3K4. However, neither are absolute requirements to form H3K4me3 sites as peaks of H3K4me3 can be formed in the absence of CpG islands and/or transcription (Barrera et al., 2008, Thomson et al., 2010).

H3K4me3 writers

Enzymes catalyzing H3K4 methylation include the MLL1, MLL2, MLL3, MLL4, SETD1A and SETD1B. The MLLs (1-4) catalyze mono- and di-methylation of H3K4 (Shilatifard, 2012, Shinsky et al., 2015). MLL1 has weak H3K4 trimethylation activity, which is stimulated in the presence of ubiquitinated H2B (Wu et al., 2013, Binda, 2013, Denissov et al., 2014, Cao et al., 2014a). SETD1A and SETD1B are responsible for most of the H3K4 trimethylation in human cells (van Nuland et al., 2013, Shinsky et al., 2015, Shilatifard, 2012).

Draft

The SETD1A and SETD1B complexes consist of SET1A or 1B, ASH2L, WDR5, RBBP5, WDR82, DPY30 and CXXC1 (Figure 1). The SETD1B complex has BOD1L, which is not present in the SETD1A complex (van Nuland et al., 2013). The prominent activity in colorectal cancer HCT116 cells is SETD1A (Salz et al., 2014b). In MDA-MB-231 triple negative breast cancer cells, both SETD1A and 1B contribute to the genesis of H3K4me3 (Salz et al., 2014a). Knocking down SETD1A reduced the cell growth and metastatic properties of colorectal cancer cells, and had a major impact on attenuating the expression of Wnt/β-catenin genes (Salz et al., 2014b). Reducing SETD1A expression also lowered the metastatic properties of triple negative breast cancer cells, MDA-MB-231 and

BT549, and attenuated the expression of matrix metalloproteinases (Salz et al., 2014a).

The recruiters of the H3K4 writers

Several players (CXXC1, uH2B, H3R2me2s) are involved in directing SETD1A/B to trimethylate H3K4 at specific sites in the 5' region of the gene body. However, we lack an appreciation of the relative role of these recruiters.

CXXC1 (a.k.a. CFP1) binds to un-methylated CpG islands. *Cxxc1* knockout murine embryonic stem cells are viable, but have several defects (e.g. inability to differentiate *in vitro*) (Tate et al., 2010). Disruption of mouse *Cxxc1* results in embryonic lethality. The knockout embryonic stem cells had lower SETD1A activity, and importantly both SETD1A and H3K4me3 were mislocalized. Similar results were obtained in CXXC1 knockdown NIH 3T3 cells (Thomson et al., 2010). These authors reported that H3K4me3 at CpG islands was greatly reduced when CXXC1 was knocked down. Expression of full length *Cxxc1* was required to restrict SETD1A and H3K4me3 to euchromatin in mammalian cells (Tate et al., 2010). The CXXC1 subunit of SETD1A/B binds to the CpG island only when it is not methylated (Xu et al., 2011).

Ubiquitinated H2B (**uH2B**) and H3 symmetrically dimethylated at R2 (**H3R2me2s**) are involved in the recruitment of the SETD1A/B complexes. We were the first to show that ubiquitination of H2B was dependent upon on-going transcription (Davie and Murphy, 1990, Nickel et al., 1989). RNF20-mediated ubiquitination of H2B was thought to be a prerequisite for trimethylation of H3K4

(Shilatifard, 2012, Ma et al., 2011). However, uH2B had no direct regulatory role on *Drosophila* SET1 (Ardehali et al., 2011). Furthermore, loss of *Drosophila* Ataxin-7, a subunit of the KAT SAGA, reduced levels of uH2B but there were no changes in H3K4me3 levels (Mohan et al., 2014). Also, *RNF20* knockdown in myoblasts resulted in a modest decrease in H3K4me3. The impact of uH2B loss on H3K4me3 at various loci was gene specific (Vethantham et al., 2012). Thus, H2B ubiquitination is not always a prerequisite for H3K4 trimethylation (Wijeweera et al., 2015).

In the mammalian genome, H3R2me2s, an active mark, colocalizes with H3K4me3 (Jahan and Davie, 2014). The PRMT catalyzing H3R2me2s is thought to be PRMT5 (a type II PRMT); however, this is not "locked down" at this time. Knockdown of this enzyme did not alter H3R2me2s levels in Rag2$^{-/-}$ Abelson-transformed pro-B cells (Yuan et al., 2012). In contrast, PRMT5 knockdown in mouse primary hepatocytes did reduce H3R2me2s at specific gene loci (about 33%) (Tsai et al., 2013). In HeLa cells, knockdown of PRMT5 and PRMT7 was effective in lowering H3R2me2s and H3K4me3 (Migliori et al., 2012). WDR5, a component of the SETD1A/1B complexes, binds to H3R2me2s, providing a mechanism by which H3R2me2s recruits SETD1A/B (Binda, 2013).

PRMT5 is also involved in pre-mRNA splicing (Bezzi et al., 2013). As inhibition of splicing reduces H3K4me3 at the 5’ splice site of the first exon (Bieberstein et al., 2012), the role of PRMT5 in splicing rather than forming H3R2me2s may be responsible for reduction in H3K4me3 observed in PRMT5 knockdown cells (Migliori et al., 2012).

https://mc06.manuscriptcentral.com/bcb-pubs

H3K4me3 erasers

There are at least four enzymes that demethylate H3K4me3, which are in the KDM5 family (5A/JARID1A/RBP2, 5B/JARID1B/PLU1, 5C/JARID1C/SMCX and 5D/JARID1D/SMCY). The expression of the KDM5 enzymes is often deregulated in cancer cells (Van Rechem and Whetstine, 2014, Blair et al., 2011). KDM5A, a DNA-binding protein, has a role in breast cancer progression and metastasis; however, this activity of KDM5A may be independent of the protein's demethylase activity (Tu et al., 2008, Cao et al., 2014b). The expression level of KDM5A correlated with the increased metastatic potential of the breast cancer cell (Cao et al., 2014b). KDM5B is up-regulated in prostate and breast cancer (Xiang et al., 2007). KDM5B is expressed at greater levels in breast cancer cells. Further, estrogen receptor positive breast cancer cells (e.g. MCF7) express greater levels of KDM5B than do triple negative (MDA-MB-231) cells (Klein et al., 2014). *KDM5B* knockdown in MDA-MB-231 cells resulted in global increases in H3K4me3 (Klein et al., 2014), while in MCF7 cells, *KDM5B* knockdown did not alter global H3K4me3 levels but did reduce H3K4me3 at specific genes (Li et al., 2011, Yamane et al., 2007). Over-expression of KDM5B in MDA-MB-231 cells suppressed the metastatic properties of these cells (Klein et al., 2014). However, knockdown of KDM5B reduces the growth of MCF7 cells (Barrett et al., 2007).

Active and poised enhancers are marked by H3K4me1, while promoters and 5' regions of transcribed genes have H3K4me3. KDM5C is recruited to both enhancers and promoters. At enhancers, KDM5C has a simulating action in

https://mc06.manuscriptcentral.com/bcb-pubs

Draft

keeping the enhancer in a H3K4me1 state. However, at promoters KDM5C has the opposite impact in dampening transcription (Outchkourov et al., 2013).

Several of the KDM5 family members are associated with transcription factors, providing a mechanism for the recruitment of these enzymes to regulatory regions of the genome. For example, KDM5B binds to the androgen receptor in prostate cancer cells (Xiang et al., 2007). KDM5B acts as a co-repressor when recruited by MYC to specific regulatory regions (Wong et al., 2012). In contrast to the recruitment of KDM5 family members to regulatory regions, the mechanism(s) by which the KDM5 enzymes are recruited to the gene body is not known. However, it is an interesting observation that all of the KDM5 family members are associated with HDAC1/2 (Li et al., 2011, Barrett et al., 2007, Tahiliani et al., 2007). KDM5B is associated with the HDAC1/2 NuRD complex (Klein et al., 2014, Li et al., 2011). It is also possible that the KDM5 family members are associated with the HDAC1/2, which is bound to RNA splicing factors (such as SRSF1 and SF3B1-4) (Khan et al., 2014). In an analysis of the HDAC1/2 complexes in mammalian cells, we observed that the unphosphorylated HDAC1/2 were associated with the pre-spliceosome E/A complex (Khan et al., 2014) (Figure 1). Possibly one or more of the KDM5 family members are recruited to the HDAC1/2 - pre-spliceosome complex and positioned to catalyze the removal of the H3K4me3 mark. Further, the KDM5 family members undergo numerous PTMs including acetylation, ubiquitination and phosphorylation (Hornbeck et al., 2015), and these PTMs could regulate activity and/or location. For example, K63-linked ubiquitination of KDM5B is

Draft

required for KDM5B to be active. SKP2 has a role in regulating KDM5B in prostate cancer cells by regulating the levels of TRAF6 which ubiquitinates KDM5B. Thus in prostate cancer cells, SKP2 maintains a low level of TRAF6 which in turn results in reduced KDM5B activity, leading to an increased level of H3K4me3 (Lu et al., 2015). The authors of this study suggest that a combinatory approach of inhibiting SKP2 and KDM5B may be an effective therapy in treating castration resistant prostate cancer (Lu et al., 2015).

H3K4me3 dynamics

Once established, H3K4me3 is a relatively stable mark in interphase and mitotic cells. The H3K4me3 mark remains associated with mitotic chromosomes acting presumably as a mitotic bookmark for expression of specific genes in the daughter cell (Terrenoire et al., 2010, Delcuve et al., 2009, Zaidi et al., 2014, Zaidi et al., 2011). However, in interphase cells the H3K4me3 mark will increase and decrease with transcription induction and cessation, respectively, of inducible genes (Edmunds et al., 2008). Primary response genes (e.g. immediate-early genes) have high basal levels of H3K4me3 at the 5' end of the gene (Hargreaves et al., 2009). Induction of immediate-early genes in quiescent mouse fibroblasts, for example, results in H3K4me3 marked nucleosomes located at the 5' end of the gene body acquiring greater levels of H3K4me3 (Edmunds et al., 2008, Hargreaves et al., 2009). However, there is a lack of information about the H3K4 methyltransferase(s) catalyzing the increase in H3K4me3. Further there is a lack of information as to the longevity of the

Draft

increased H3K4me3 following the onset of the induction. It is evident that the increased level of H3K4me3 along the gene body of immediate-early genes is sustained well after the expression of the induced immediate-early gene ceases (Edmunds et al., 2008). It would be expected that the levels of H3K4me3 at the 5' region of induced genes in mammalian cells would decline over time to a basal level following the cessation of transcription, and that this process would be catalyzed by one or more of the KDM5 family members (Ng et al., 2003, Riffo-Campos et al., 2015, Edmunds et al., 2008). However, this has yet to be reported. Studies in yeast show that the H3K4me3 mark persists well after the induced gene's transcription ceases and it takes some time (up to 5 hours) for H3K4me3 levels to go back to basal levels (Ng et al., 2003). Thus demethylation of H3K4me3 appears to be a slow process.

Draft

H3K4me3 readers

H3K4me3 has several readers including CHD1, FACT, the PAF complex, U2 snRNP and U2AF65, SAGA-associated factor 29, and the ING proteins (Guerillon et al., 2013, Buratowski and Kim, 2010, Shi et al., 2006, Schram et al., 2013, Sims et al., 2007) (Figure 1). CHD1 is an ATP-dependent chromatin remodeler, which binds to the SF3A subunit of the U2 snRNP complex. ING3, ING4/5 and ING5 are associated with the MYST family of KATs, including Tip60, HBO1 and MOZ/MORF (Lalonde et al., 2013). FACT, consisting of Spt16 and SSRP1, has a role in the removal and the re-depositing of the H2A-H2B dimer during the passage of RNA polymerase II (Kwak and Lis, 2013). The PAF

complex has a role in transcriptional elongation. The PAF complex is required by RNF20 to ubiquitinate H2B (Kwak and Lis, 2013). The U2 snRNP complex recognizes the 3' splice site in an early step in splicing (Figure 1). The U2 snRNP has SF3A1-A3 and SF3B1-B5. U2AF65 is a U2 snRNP auxiliary factor that is an essential component of the pre-mRNA splicing machinery (Prigge et al., 2009). *Clearly H3K4me3 is a nexus for proteins involved in transcriptional elongation, chromatin remodeling and modification, and pre-mRNA splicing* (Figure 1).

H3K4me3 marked nucleosomes may also prevent the binding of epigenetic modifiers. For example, DNA methyltransferases, DNMT3A, DNMT3A, and co-factor DNMT3L, all contain an ADD (ATRX-DNMT3-DNMT3L) domain, which binds to H3K9 methylated and unmodified histones via their PHD domain within ADD. H3K4me3 prevents the binding of the ADD domain and these DNA methyltransferases to the H3 tail (Rose and Klose, 2014).

Draft

RNA and epigenetic modifiers

Nuclear RNA, coding or non-coding (**ncRNA**), regulates the structure and function of chromatin (Gardini and Shiekhattar, 2015, Guil and Esteller, 2012, Caudron-Herger et al., 2011, Caudron-Herger and Rippe, 2012, Natoli and Andrau, 2012, Rodriguez-Campos and Azorin, 2007, Orom and Shiekhattar, 2011, Ansari and Morse, 2013, Lai et al., 2013). There is mounting evidence that RNA-bound chromatin modifiers and remodelers play key roles in determining histone PTM positions along chromatin. RNA-bound chromatin modifiers, architectural proteins, chromatin-associated proteins and transcriptional

regulators include KAT2B, HDAC1/2, SWI/SNF, CTCF, MeCP2, WDR5/MLL (an H3K4 methyltransferase), EZH2 (a repressor), and mediator (Khan et al., 2014, Sjolinder et al., 2005, Tyagi et al., 2009, Guil et al., 2012, Khan et al., 2012, Maxwell et al., 2013, Tsai et al., 2010, Guil and Esteller, 2012, Gardini and Shiekhattar, 2015). SETD1B binds to the *MPL* transcript (Xiao et al., 2014). Thus, there is evidence that the enzymes catalyzing H3K4 trimethylation are associated with RNA. The association of these chromatin modifiers with RNA is often indirect, with the factor being recruited to RNA by RNA-binding proteins. For example, KAT2B is recruited to RNA by hnRNP U (Obrdlik et al., 2008). Several of the epigenetic modifiers associating with RNA binding proteins have roles in pre-mRNA splicing (e.g. KAT2B, HDAC1/2, MeCP2) (Khan et al., 2012, Khan et al., 2014, Liyanage et al., 2014) (Figure 1).

Draft

We recently reported that HDAC1 and HDAC2 were recruited to the gene body of a transcribed gene, with the recruitment of the HDACs being transcription-dependent (Khan et al., 2014). However, the HDAC1 and 2 were not associated with the chromatin but with the transcript. Using a dual cross-linking ChIP assay, HDAC1 and 2 appeared to be bound to the gene body of the transcribed gene. Yet, when the fragmented chromatin was digested with RNase A before adding the anti-HDAC1 or 2 antibodies, the association of the HDACs with the transcribed gene body was lost. These studies highlighted a caution in interpreting ChIP assay data. RNA-binding proteins involved in pre-mRNA splicing will be cross-linked to DNA with formaldehyde and will appear to be associated with the coding region of transcribed genes (Sapra et al., 2009).

However, addition of an RNase digestion step before doing the immunoprecipitation will demonstrate whether the RNA-binding protein is bound to RNA and not to chromatin. To demonstrate that HDAC1/2 and KAT2B were bound to the *MCL1* transcript, we used a dual cross-linking CLIP protocol (Khan and Davie, 2014, Khan et al., 2014). As KDM5 family members are associated with HDAC1/2, it is possible that the KDM5 enzymes are also associated with RNA. Experimental approaches such as those described above will be required to determine whether this idea is correct.

H3K4me3 and pre-mRNA splicing efficiency

Reducing H3K4me3 levels by knocking down of *ASH2L*, a component of the SETD1A/B enzyme complex, lowered the association of U2 snRNP with chromatin and attenuated pre-mRNA splicing efficiency of an interferon inducible gene (Sims et al., 2007). WDR5, ASH2L and RBBP5 form a core complex called WAR that is in SETD1A/1B (Figure 1). Knockdown of *WDR5* or *ASH2L* reduced the efficiency of immediate early gene (*FOS*, *ZFP36*) transcript processing (Teoh and Sharrocks, 2014). H3K4me3 at the *FOS* promoter was unaffected in cells with *RBBP5* knocked down. The authors concluded that H3K4me3 at the *FOS* promoter did have a role in the processing of the *FOS* transcript. It should be noted that these authors only studied the H3K4me3 at the region upstream of the TSS and did not investigate the status of H3K4me3 located at the 5' region of the gene body (Teoh and Sharrocks, 2014). We speculate that the H3K4me3 peak downstream of the *FOS* TSS would be reduced in cells in which any of the WAR

Draft

complex proteins were knocked down and that this decrease in H3K4me3 would correlate with the increased levels of unspliced *FOS* transcripts.

Dynamic histone acetylation, H3K4me3 and alternative pre-mRNA splicing

Nucleosomes with H3K4me3 are selectively dynamically acetylated (Wang et al., 2009, Hendzel and Davie, 1991, Hazzalin and Mahadevan, 2005). With the inhibition of HDACs, H3K4me3 modified nucleosomes become hyperacetylated. Several KAT complexes, including KAT2B and KAT7, have among their associated proteins those that read the H3K4me3 mark (SAGA-associated factor 29, and the ING proteins). In situations where the H3K4me3 modified nucleosome is positioned over an alternative exon, the incubation of cells with HDAC inhibitors may result in changes in alternative splicing of the pre-mRNA. For example, the *MCL1* gene in HCT116 cells has H3K4me3 modified nucleosomes positioned at the end of exon 1 and over exon 2. The *MCL1* gene undergoes alternative splicing of exon 2 and produces a protein that either prevents or supports cell death. The long form MCL1L, in which the transcript retains exon 2 is an anti-apoptotic protein, while the short form MCL1S, in which the transcript has lost exon 2 is pro-apoptotic (Khan et al., 2014). In cancer cells the anti-apoptotic form MCL1L predominates. A brief 2 hour treatment of the HCT116 cells HDAC inhibitors results in the increased exclusion of exon 2 during pre-mRNA splicing. The hyperacetylated state of the alternative exon 2 nucleosome results in increased rates of elongation through this nucleosome, preventing the association of key splicing factors (e.g. SRSF1) with the pre-

Draft

mRNA. These events result in exon 2 exclusion (Khan et al., 2014). This study demonstrates how the strategic H3K4me3 modification of a nucleosome in a gene body confers to this nucleosome unique properties enabling a response to environmental changes. Thus in the context of a signaling module, the H3K4me3 modified nucleosome is quite different from its neighboring nucleosomes.

Disease and H3K4me3 nexus

Proteins/enzymes that regulate the level and location of H3K4me3 are deregulated in breast cancer. Further H3K4me3-binding proteins (readers) with roles in chromatin modification, elongation and RNA splicing are mutated in breast cancer. According to the Cancer Genome Atlas, KDM5B is frequently over-expressed in invasive ductal carcinoma cancer (Lu et al., 1999, Cancer Genome Atlas, 2012, Rasmussen and Staller, 2014, Blair et al., 2011). KDM5C is frequently mutated in X-linked mental retardation (Tzschach et al., 2006, Blair et al., 2011, Jensen et al., 2010). KDM5C is also mutated in clear cell renal cell carcinomas (Blair et al., 2011). *ING1, 2* & *4* genes are often deleted in breast cancer (Guerillon et al., 2013). The ING proteins are components of several epigenetic modifiers and readers of the H3K4me3 mark. *SF3B1* and *U2AF1* genes are often mutated in disease states (Scott and Rebel, 2013). *SF3B1* and *SF3B3* genes are mutated in breast cancer (Scott and Rebel, 2013, Cancer Genome Atlas, 2012, Maguire et al., 2015). SF3B1 and SF3B3 are part of the H3K4me3-binding U2 snRNP, which regulates splicing. A picture is emerging that H3K4me3 is the hub for several events, which regulate splicing and

Draft

chromatin dynamics, contributing to altered gene expression in breast cancer (Figure 1).

The genes coding for the MLL enzymes are often mutated in disease states. Mutations in MLL2 and MLL3 are found in the Kabuki 1 and Kleefstra syndromes, respectively (Parkel et al., 2013). Further cancer mutations in MLL3 result in an enzyme catalyzing H3K4me3, which is in contrast to the wild type enzyme that produces H3K4me1 (Weirich et al., 2015). Fusion proteins (MLL-AF9, MLL-ENL) produced by translocation events in cancer cells, result in mistargeting of the H3K4 methyltransferase activity of the MLL enzyme (Yokoyama, 2015, Li and Ernst, 2014).

There are many examples of how cancer mutations in epigenetic modifiers such as EZH2 alter enzymatic activity. Recently, Weirich et al reported that the mutation Y4848S in MLL3 converts MLL3 from a H3K4 monomethyltransferase to a trimethyltransferase, resulting in increased levels of H3K4me3 in cancer cells (Weirich et al., 2015). As MLL3/4 monomethylate H3K4 at enhancers (Hu et al., 2013), such a mutation in MLL3 may alter the activity of enhancers (Weirich et al., 2015).

Draft

Inhibitors of KDM5

Towards understanding the biological role of KDM5 family members and importantly in the treatment of cancers over-expressing KDM5 family members, several KDM5 inhibitors have been synthesized (Sayegh et al., 2013, Thinnes et al., 2014). Inhibitors of KDM5B include 2-4(4-methylphenyl)-1,2-benzisothiazol-

3(2H)-one (PBIT), which inhibits KDM5B with an IC_{50} of 3 μM *in vitro* (Sayegh et al., 2013). KDM5A and KDM5C but not KDM6A or KDM6B are inhibited by PBIT (Thinnes et al., 2014). PBIT is derived from a commonly used microbicide and fungicide, benzisothiazolinone, used in home cleaning products (Sayegh et al., 2013). PBIT prevented the proliferation of cancer cells (UACC-812 breast cancer cells) over-expressing KDM5B. PBIT treatment of estrogen receptor positive MCF7 breast cancer cells, which have a greater expression of KDM5B than triple negative MDA-MB-231 breast cancer cells, significantly increased H3K4me3 levels in these cells. The PBIT analogue ebselen is somewhat less potent in inhibiting KDM5B than PBIT, with ebselen having an IC_{50} of about 6 μM.

Draft

Inhibitors of pre-mRNA splicing and impact on H3K4me3

There are several inhibitors of pre-mRNA splicing. FR901464, meayamycin A/B and spliceostatin A bind to the splicing factor 3b complex (specifically SF3B1) and inhibit pre-mRNA splicing (Albert et al., 2009, Gao and Koide, 2013, Bieberstein et al., 2012, Kaida et al., 2007). Meayamycin A is a potent inhibitor of cancer cell (e.g. MCF7, MDA-MB-231, HCT116) proliferation (Albert et al., 2009). The alternative splicing of *MCL1* and *BCL2L1* (also known as *BCL-XL*) is regulated by SRSF1 and SF3B1 (Moore et al., 2010). Interestingly HDAC1 and 2 interact with both of these splicing factors (Khan et al., 2014). Treatment of cancer cells (H1299, A549) with meayamycin B results in an increase in the production of *MCL1S*, the transcript lacking exon 2 and coding for an pro-apoptotic protein (Gao and Koide, 2013). Treatment of HeLa cells with

spliceostatin A resulted in the loss of H3K4me3 peak positioned at the end of exon 1 of a *FOS* construct (Bieberstein et al., 2012). The mechanism by which inhibition of splicing alters H3K4me3 is not known. However, it is possible that recruitment of the SETD1A/B bound to RNA is disturbed with the splicing inhibitors. Thus H3K4me3 nucleosome signaling impacts splicing and vice versa splicing impacts H3K4me3.

Concluding remarks

H3K4me3 is one of the most studied marks used in identifying promoter regions in epigenome mapping studies (Roadmap Epigenomics et al., 2015). Although often stated as a mark of active/poised promoters, the major H3K4me3 peak in mammalian cells resides at the end of exon 1 at the 5' splice site. In establishing the H3K4me3 mark downstream of the TSS, CpG islands appear to have a major role in recruiting the SETD1A/B complex, an event that is not dependent on transcription. However, the H3K4me3 peaks located upstream and downstream of the TSS may be catalyzed by different H3K4me3 KMTs, and this will need to be resolved by ChIP assays. For genes such as *MCL1* there are multiple peaks of H3K4me3 downstream of exon 1 that are in similar positions in different cell types. The processes involved in forming these peaks are not known. Possibly chromatin/nucleosome organization or the location of the enzymes and/or recruiters generating the H3K4me3 mark are involved. Once the H3K4me3 has been established, it appears to be a "stable" epigenetic mark that remains associated with mitotic chromosomes. However, we do not know if the

https://mc06.manuscriptcentral.com/bcb-pubs

genomic location of H3K4me3 on mitotic chromosome is the same as in interphase chromatin, which again calls upon ChIP sequencing studies to sort this out.

Studies on inducible genes demonstrate that the level but not position of the H3K4me3 modification can be altered. The levels of H3K4me3 may rise and fall but the H3K4me3 mark is not erased. We have much to learn about the enzymes involved in increasing (and decreasing) H3K4me3 levels at sites in the body of inducible genes (e.g. immediate-early genes). Also the mechanisms by which the H3K4me3 KMTs are recruited to inducible genes need to be determined. It is conceivable that the H3K4me3 KMT is associated with the transcript, but this remains to be demonstrated.

Draft

Current evidence shows that pre-mRNA splicing, which occurs co-transcriptionally in mammalian cells, has a role in regulating H3K4me3 levels at sites downstream of the TSS. Further, the enzymes involved in H3K4 trimethylation have a role in splicing efficiency. Future studies will be required to identify the multiple players involved in fine-tuning H3K4me3 levels and splicing events.

In closing, H3K4me3 dynamics is a great example of the interplay between transcription and pre-mRNA splicing, producing a nucleosome that is singled out to be a strong signalling module for recruitment of chromatin remodelers and lysine acetyltransferases.

Acknowledgements

This work was supported by a grant from the Canadian Breast Cancer Foundation and a Canada Research Chair to JRD. The authors acknowledge the strong support of the Children's Hospital Research Institute of Manitoba and CancerCare Manitoba Foundation.

References

ALBERT, B. J., MCPHERSON, P. A., O'BRIEN, K., CZAICKI, N. L., DESTEFINO, V., OSMAN, S., LI, M., DAY, B. W., GRABOWSKI, P. J., MOORE, M. J., VOGT, A. & KOIDE, K. 2009. Meayamycin inhibits pre-messenger RNA splicing and exhibits picomolar activity against multidrug-resistant cells. *Mol Cancer Ther,* 8**,** 2308-18.

ALLIS, C. D., BERGER, S. L., COTE, J., DENT, S., JENUWIEN, T., KOUZARIDES, T., PILLUS, L., REINBERG, D., SHI, Y., SHIEKHATTAR, R., SHILATIFARD, A., WORKMAN, J. & ZHANG, Y. 2007. New nomenclature for chromatin-modifying enzymes. *Cell,* 131**,** 633-6.

ANSARI, S. A. & MORSE, R. H. 2013. Mechanisms of Mediator complex action in transcriptional activation. *Cell Mol Life Sci,* 70**,** 2743-56.

ARDEHALI, M. B., MEI, A., ZOBECK, K. L., CARON, M., LIS, J. T. & KUSCH, T. 2011. Drosophila Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription. *EMBO J,* 30**,** 2817-28.

BARRERA, L. O., LI, Z., SMITH, A. D., ARDEN, K. C., CAVENEE, W. K., ZHANG, M. Q., GREEN, R. D. & REN, B. 2008. Genome-wide mapping and analysis of active promoters in mouse embryonic stem cells and adult organs. *Genome Res,* 18**,** 46-59.

BARRETT, A., SANTANGELO, S., TAN, K., CATCHPOLE, S., ROBERTS, K., SPENCER-DENE, B., HALL, D., SCIBETTA, A., BURCHELL, J., VERDIN, E., FREEMONT, P. & TAYLOR-PAPADIMITRIOU, J. 2007. Breast cancer associated transcriptional repressor PLU-1/JARID1B interacts directly with histone deacetylases. *Int J Cancer,* 121**,** 265-75.

BARSKI, A., CUDDAPAH, S., CUI, K., ROH, T. Y., SCHONES, D. E., WANG, Z., WEI, G., CHEPELEV, I. & ZHAO, K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell,* 129**,** 823-37.

BENTLEY, D. L. 2014. Coupling mRNA processing with transcription in time and space. *Nat Rev Genet,* 15**,** 163-75.

BEZZI, M., TEO, S. X., MULLER, J., MOK, W. C., SAHU, S. K., VARDY, L. A., BONDAY, Z. Q. & GUCCIONE, E. 2013. Regulation of constitutive and alternative splicing by PRMT5 reveals a role for Mdm4 pre-mRNA in sensing defects in the spliceosomal machinery. *Genes Dev,* 27**,** 1903-16.

BIEBERSTEIN, N. I., CARRILLO OESTERREICH, F., STRAUBE, K. & NEUGEBAUER, K. M. 2012. First exon length controls active chromatin signatures and transcription. *Cell Rep,* 2**,** 62-8.
BINDA, O. 2013. On your histone mark, SET, methylate! *Epigenetics,* 8**,** 457-63.
BLAIR, L. P., CAO, J., ZOU, M. R., SAYEGH, J. & YAN, Q. 2011. Epigenetic Regulation by Lysine Demethylase 5 (KDM5) Enzymes in Cancer. *Cancers (Basel),* 3**,** 1383-404.
BRAUNSCHWEIG, U., GUEROUSSOV, S., PLOCIK, A. M., GRAVELEY, B. R. & BLENCOWE, B. J. 2013. Dynamic integration of splicing within gene regulatory pathways. *Cell,* 152**,** 1252-69.
BURATOWSKI, S. & KIM, T. 2010. The role of cotranscriptional histone methylations. *Cold Spring Harb Symp Quant Biol,* 75**,** 95-102.
CANCER GENOME ATLAS, N. 2012. Comprehensive molecular portraits of human breast tumours. *Nature,* 490**,** 61-70.
CAO, F., TOWNSEND, E. C., KARATAS, H., XU, J., LI, L., LEE, S., LIU, L., CHEN, Y., OUILLETTE, P., ZHU, J., HESS, J. L., ATADJA, P., LEI, M., QIN, Z. S., MALEK, S., WANG, S. & DOU, Y. 2014a. Targeting MLL1 H3K4 methyltransferase activity in mixed-lineage leukemia. *Mol Cell,* 53**,** 247-61.
CAO, J., LIU, Z., CHEUNG, W. K., ZHAO, M., CHEN, S. Y., CHAN, S. W., BOOTH, C. J., NGUYEN, D. X. & YAN, Q. 2014b. Histone demethylase RBP2 is critical for breast cancer progression and metastasis. *Cell Rep,* 6**,** 868-77.
CAUDRON-HERGER, M., MULLER-OTT, K., MALLM, J. P., MARTH, C., SCHMIDT, U., FEJES-TOTH, K. & RIPPE, K. 2011. Coding RNAs with a non-coding function: maintenance of open chromatin structure. *Nucleus,* 2**,** 410-24.
CAUDRON-HERGER, M. & RIPPE, K. 2012. Nuclear architecture by RNA. *Curr Opin Genet Dev,* 22**,** 179-87.
CHEN, R. A., STEMPOR, P., DOWN, T. A., ZEISER, E., FEUER, S. K. & AHRINGER, J. 2014. Extreme HOT regions are CpG-dense promoters in C. elegans and humans. *Genome Res,* 24**,** 1138-46.
CHODAVARAPU, R. K., FENG, S., BERNATAVICHUTE, Y. V., CHEN, P. Y., STROUD, H., YU, Y., HETZEL, J. A., KUO, F., KIM, J., COKUS, S. J., CASERO, D., BERNAL, M., HUIJSER, P., CLARK, A. T., KRAMER, U., MERCHANT, S. S., ZHANG, X., JACOBSEN, S. E. & PELLEGRINI, M. 2010. Relationship between nucleosome positioning and DNA methylation. *Nature,* 466**,** 388-92.
DAVIE, J. R. & MURPHY, L. C. 1990. Level of ubiquitinated histone H2B in chromatin is coupled to ongoing transcription. *Biochemistry,* 29**,** 4752-7.
DE ALMEIDA, S. F. & CARMO-FONSECA, M. 2014. Reciprocal regulatory links between cotranscriptional splicing and chromatin. *Semin Cell Dev Biol,* 32**,** 2-10.
DEATON, A. M., WEBB, S., KERR, A. R., ILLINGWORTH, R. S., GUY, J., ANDREWS, R. & BIRD, A. 2011. Cell type-specific DNA methylation at intragenic CpG islands in the immune system. *Genome Res,* 21**,** 1074-86.

Draft

DELCUVE, G. P., RASTEGAR, M. & DAVIE, J. R. 2009. Epigenetic control. *J Cell Physiol,* 219**,** 243-50.
DENISSOV, S., HOFEMEISTER, H., MARKS, H., KRANZ, A., CIOTTA, G., SINGH, S., ANASTASSIADIS, K., STUNNENBERG, H. G. & STEWART, A. F. 2014. Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. *Development,* 141**,** 526-37.
DUJARDIN, G., LAFAILLE, C., PETRILLO, E., BUGGIANO, V., GOMEZ ACUNA, L. I., FISZBEIN, A., GODOY HERZ, M. A., NIETO MORENO, N., MUNOZ, M. J., ALLO, M., SCHOR, I. E. & KORNBLIHTT, A. R. 2013. Transcriptional elongation and alternative splicing. *Biochim Biophys Acta,* 1829**,** 134-40.
EDMUNDS, J. W., MAHADEVAN, L. C. & CLAYTON, A. L. 2008. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J,* 27**,** 406-20.
ENCODEPROJECTCONSORTIUM 2011. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol,* 9**,** e1001046.
ENCODEPROJECTCONSORTIUM 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature,* 489**,** 57-74.
GAO, Y. & KOIDE, K. 2013. Chemical perturbation of Mcl-1 pre-mRNA splicing to induce apoptosis in cancer cells. *ACS Chem Biol,* 8**,** 895-900.
GARDINI, A. & SHIEKHATTAR, R. 2015. The many faces of long noncoding RNAs. *FEBS J,* 282**,** 1647-1657.
GUENTHER, M. G., LEVINE, S. S., BOYER, L. A., JAENISCH, R. & YOUNG, R. A. 2007. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell,* 130**,** 77-88.
GUERILLON, C., LARRIEU, D. & PEDEUX, R. 2013. ING1 and ING2: multifaceted tumor suppressor genes. *Cell Mol Life Sci,* 70**,** 3753-72.
GUIL, S. & ESTELLER, M. 2012. Cis-acting noncoding RNAs: friends and foes. *Nat Struct Mol Biol,* 19**,** 1068-75.
GUIL, S., SOLER, M., PORTELA, A., CARRERE, J., FONALLERAS, E., GOMEZ, A., VILLANUEVA, A. & ESTELLER, M. 2012. Intronic RNAs mediate EZH2 regulation of epigenetic targets. *Nat Struct Mol Biol,* 19**,** 664-70.
HARGREAVES, D. C., HORNG, T. & MEDZHITOV, R. 2009. Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell,* 138**,** 129-45.
HAZZALIN, C. A. & MAHADEVAN, L. C. 2005. Dynamic acetylation of all lysine 4-methylated histone H3 in the mouse nucleus: analysis at c-fos and c-jun. *PLoS Biol,* 3**,** e393.
HENDZEL, M. J. & DAVIE, J. R. 1991. Dynamically acetylated histones of chicken erythrocytes are selectively methylated. *Biochem J,* 273 (Pt 3)**,** 753-8.
HNILICOVA, J. & STANEK, D. 2011. Where splicing joins chromatin. *Nucleus,* 2**,** 182-8.

Draft

HORNBECK, P. V., ZHANG, B., MURRAY, B., KORNHAUSER, J. M., LATHAM, V. & SKRZYPEK, E. 2015. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res,* 43**,** D512-20.

HU, D., GAO, X., MORGAN, M. A., HERZ, H. M., SMITH, E. R. & SHILATIFARD, A. 2013. The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. *Mol Cell Biol,* 33**,** 4745-54.

IP, J. Y., SCHMIDT, D., PAN, Q., RAMANI, A. K., FRASER, A. G., ODOM, D. T. & BLENCOWE, B. J. 2011. Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. *Genome Res,* 21**,** 390-401.

JAHAN, S. & DAVIE, J. R. 2014. Protein arginine methyltransferases (PRMTs): Role in chromatin organization. *Adv Biol Regul*.

JENSEN, L. R., BARTENSCHLAGER, H., RUJIRABANJERD, S., TZSCHACH, A., NUMANN, A., JANECKE, A. R., SPORLE, R., STRICKER, S., RAYNAUD, M., NELSON, J., HACKETT, A., FRYNS, J. P., CHELLY, J., DE BROUWER, A. P., HAMEL, B., GECZ, J., ROPERS, H. H. & KUSS, A. W. 2010. A distinctive gene expression fingerprint in mentally retarded male patients reflects disease-causing defects in the histone demethylase KDM5C. *Pathogenetics,* 3**,** 2.

KAIDA, D., MOTOYOSHI, H., TASHIRO, E., NOJIMA, T., HAGIWARA, M., ISHIGAMI, K., WATANABE, H., KITAHARA, T., YOSHIDA, T., NAKAJIMA, H., TANI, T., HORINOUCHI, S. & YOSHIDA, M. 2007. Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat Chem Biol,* 3**,** 576-83.

KALSOTRA, A. & COOPER, T. A. 2011. Functional consequences of developmentally regulated alternative splicing. *Nat Rev Genet,* 12**,** 715-29.

KHAN, D. H. & DAVIE, J. R. 2014. Dual cross-linking ribonucleoprotein immunoprecipitation assay. *Biochem Cell Biol,* 92**,** 317-9.

KHAN, D. H., GONZALEZ, C., COOPER, C., SUN, J. M., CHEN, H. Y., HEALY, S., XU, W., SMITH, K. T., WORKMAN, J. L., LEYGUE, E. & DAVIE, J. R. 2014. RNA-dependent dynamic histone acetylation regulates MCL1 alternative splicing. *Nucleic Acids Res,* 42**,** 1656-70.

KHAN, D. H., JAHAN, S. & DAVIE, J. R. 2012. Pre-mRNA splicing: role of epigenetics and implications in disease. *Adv Biol Regul,* 52**,** 377-88.

KLEIN, B. J., PIAO, L., XI, Y., RINCON-ARANO, H., ROTHBART, S. B., PENG, D., WEN, H., LARSON, C., ZHANG, X., ZHENG, X., CORTAZAR, M. A., PENA, P. V., MANGAN, A., BENTLEY, D. L., STRAHL, B. D., GROUDINE, M., LI, W., SHI, X. & KUTATELADZE, T. G. 2014. The histone-H3K4-specific demethylase KDM5B binds to its substrate and product through distinct PHD fingers. *Cell Rep,* 6**,** 325-35.

KORNBLIHTT, A. R., SCHOR, I. E., ALLO, M., DUJARDIN, G., PETRILLO, E. & MUNOZ, M. J. 2013. Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nat Rev Mol Cell Biol,* 14**,** 153-65.

Draft

KWAK, H., FUDA, N. J., CORE, L. J. & LIS, J. T. 2013. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science,* 339**,** 950-3.
KWAK, H. & LIS, J. T. 2013. Control of transcriptional elongation. *Annu Rev Genet,* 47**,** 483-508.
LAI, F., OROM, U. A., CESARONI, M., BERINGER, M., TAATJES, D. J., BLOBEL, G. A. & SHIEKHATTAR, R. 2013. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature,* 494**,** 497-501.
LALONDE, M. E., AVVAKUMOV, N., GLASS, K. C., JONCAS, F. H., SAKSOUK, N., HOLLIDAY, M., PAQUET, E., YAN, K., TONG, Q., KLEIN, B. J., TAN, S., YANG, X. J., KUTATELADZE, T. G. & COTE, J. 2013. Exchange of associated factors directs a switch in HBO1 acetyltransferase histone tail specificity. *Genes Dev,* 27**,** 2009-24.
LI, B. E. & ERNST, P. 2014. Two decades of leukemia oncoprotein epistasis: the MLL1 paradigm for epigenetic deregulation in leukemia. *Exp Hematol,* 42**,** 995-1012.
LI, Q., SHI, L., GUI, B., YU, W., WANG, J., ZHANG, D., HAN, X., YAO, Z. & SHANG, Y. 2011. Binding of the JmjC demethylase JARID1B to LSD1/NuRD suppresses angiogenesis and metastasis in breast cancer cells by repressing chemokine CCL14. *Cancer Res,* 71**,** 6899-908.
LIYANAGE, V. R., JARMASZ, J. S., MURUGESHAN, N., DEL BIGIO, M. R., RASTEGAR, M. & DAVIE, J. R. 2014. DNA Modifications: Function and Applications in Normal and Disease States. *Biology (Basel),* 3**,** 670-723.
LU, P. J., SUNDQUIST, K., BAECKSTROM, D., POULSOM, R., HANBY, A., MEIER-EWERT, S., JONES, T., MITCHELL, M., PITHA-ROWE, P., FREEMONT, P. & TAYLOR-PAPADIMITRIOU, J. 1999. A novel gene (PLU-1) containing highly conserved putative DNA/chromatin binding motifs is specifically up-regulated in breast cancer. *J Biol Chem,* 274**,** 15633-45.
LU, W., LIU, S., LI, B., XIE, Y., ADHIAMBO, C., YANG, Q., BALLARD, B. R., NAKAYAMA, K. I., MATUSIK, R. J. & CHEN, Z. 2015. SKP2 inactivation suppresses prostate tumorigenesis by mediating JARID1B ubiquitination. *Oncotarget,* 6**,** 771-88.
LUCO, R. F., PAN, Q., TOMINAGA, K., BLENCOWE, B. J., PEREIRA-SMITH, O. M. & MISTELI, T. 2010. Regulation of alternative splicing by histone modifications. *Science,* 327**,** 996-1000.
MA, M. K., HEATH, C., HAIR, A. & WEST, A. G. 2011. Histone crosstalk directed by H2B ubiquitination is required for chromatin boundary integrity. *PLoS Genet,* 7**,** e1002175.
MAGUIRE, S. L., LEONIDOU, A., WAI, P., MARCHIO, C., NG, C. K., SAPINO, A., SALOMON, A. V., REIS-FILHO, J. S., WEIGELT, B. & NATRAJAN, R. C. 2015. SF3B1 mutations constitute a novel therapeutic target in breast cancer. *J Pathol,* 235**,** 571-80.
MAUNAKEA, A. K., NAGARAJAN, R. P., BILENKY, M., BALLINGER, T. J., D'SOUZA, C., FOUSE, S. D., JOHNSON, B. E., HONG, C., NIELSEN, C.,

ZHAO, Y., TURECKI, G., DELANEY, A., VARHOL, R., THIESSEN, N., SHCHORS, K., HEINE, V. M., ROWITCH, D. H., XING, X., FIORE, C., SCHILLEBEECKX, M., JONES, S. J., HAUSSLER, D., MARRA, M. A., HIRST, M., WANG, T. & COSTELLO, J. F. 2010. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature,* 466**,** 253-7.

MAXWELL, S. S., PELKA, G. J., TAM, P. P. & EL-OSTA, A. 2013. Chromatin context and ncRNA highlight targets of MeCP2 in brain. *RNA Biol,* 10**,** 1741-57.

MAZE, I., NOH, K. M., SOSHNEV, A. A. & ALLIS, C. D. 2014. Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nat Rev Genet,* 15**,** 259-71.

MIGLIORI, V., MULLER, J., PHALKE, S., LOW, D., BEZZI, M., MOK, W. C., SAHU, S. K., GUNARATNE, J., CAPASSO, P., BASSI, C., CECATIELLO, V., DE MARCO, A., BLACKSTOCK, W., KUZNETSOV, V., AMATI, B., MAPELLI, M. & GUCCIONE, E. 2012. Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance. *Nat Struct Mol Biol,* 19**,** 136-44.

MOHAN, R. D., DIALYNAS, G., WEAKE, V. M., LIU, J., MARTIN-BROWN, S., FLORENS, L., WASHBURN, M. P., WORKMAN, J. L. & ABMAYR, S. M. 2014. Loss of Drosophila Ataxin-7, a SAGA subunit, reduces H2B ubiquitination and leads to neural and retinal degeneration. *Genes Dev,* 28**,** 259-72.

MOORE, M. J., WANG, Q., KENNEDY, C. J. & SILVER, P. A. 2010. An alternative splicing network links cell-cycle control to apoptosis. *Cell,* 142**,** 625-36.

NATOLI, G. & ANDRAU, J. C. 2012. Noncoding transcription at enhancers: general principles and functional models. *Annu Rev Genet,* 46**,** 1-19.

NG, H. H., ROBERT, F., YOUNG, R. A. & STRUHL, K. 2003. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell,* 11**,** 709-19.

NICKEL, B. E., ALLIS, C. D. & DAVIE, J. R. 1989. Ubiquitinated histone H2B is preferentially located in transcriptionally active chromatin. *Biochemistry,* 28**,** 958-63.

OBRDLIK, A., KUKALEV, A., LOUVET, E., FARRANTS, A. K., CAPUTO, L. & PERCIPALLE, P. 2008. The histone acetyltransferase PCAF associates with actin and hnRNP U for RNA polymerase II transcription. *Mol Cell Biol,* 28**,** 6342-57.

OROM, U. A. & SHIEKHATTAR, R. 2011. Noncoding RNAs and enhancers: complications of a long-distance relationship. *Trends Genet,* 27**,** 433-9.

OUTCHKOUROV, N. S., MUINO, J. M., KAUFMANN, K., VAN IJCKEN, W. F., GROOT KOERKAMP, M. J., VAN LEENEN, D., DE GRAAF, P., HOLSTEGE, F. C., GROSVELD, F. G. & TIMMERS, H. T. 2013. Balancing of histone H3K4 methylation states by the Kdm5c/SMCX

Draft

histone demethylase modulates promoter and enhancer function. *Cell Rep,* 3**,** 1071-9.
PAL, S., GUPTA, R., KIM, H., WICKRAMASINGHE, P., BAUBET, V., SHOWE, L. C., DAHMANE, N. & DAVULURI, R. V. 2011. Alternative transcription exceeds alternative splicing in generating the transcriptome diversity of cerebellar development. *Genome Res,* 21**,** 1260-72.
PARKEL, S., LOPEZ-ATALAYA, J. P. & BARCO, A. 2013. Histone H3 lysine methylation in cognition and intellectual disability disorders. *Learn Mem,* 20**,** 570-9.
PRIGGE, J. R., IVERSON, S. V., SIDERS, A. M. & SCHMIDT, E. E. 2009. Interactome for auxiliary splicing factor U2AF(65) suggests diverse roles. *Biochim Biophys Acta,* 1789**,** 487-92.
RAHL, P. B., LIN, C. Y., SEILA, A. C., FLYNN, R. A., MCCUINE, S., BURGE, C. B., SHARP, P. A. & YOUNG, R. A. 2010. c-Myc regulates transcriptional pause release. *Cell,* 141**,** 432-45.
RASMUSSEN, P. B. & STALLER, P. 2014. The KDM5 family of histone demethylases as targets in oncology drug discovery. *Epigenomics,* 6**,** 277-86.
RIFFO-CAMPOS, A. L., CASTILLO, J., TUR, G., GONZALEZ-FIGUEROA, P., GEORGIEVA, E. I., RODRIGUEZ, J. L., LOPEZ-RODAS, G., RODRIGO, M. I. & FRANCO, L. 2015. Nucleosome-specific, Time-dependent Changes in Histone Modifications during Activation of the Early Growth Response 1 (Egr1) Gene. *J Biol Chem,* 290**,** 197-208.
ROADMAP EPIGENOMICS, C., KUNDAJE, A., MEULEMAN, W., ERNST, J., BILENKY, M., YEN, A., HERAVI-MOUSSAVI, A., KHERADPOUR, P., ZHANG, Z., WANG, J., ZILLER, M. J., AMIN, V., WHITAKER, J. W., SCHULTZ, M. D., WARD, L. D., SARKAR, A., QUON, G., SANDSTROM, R. S., EATON, M. L., WU, Y. C., PFENNING, A. R., WANG, X., CLAUSSNITZER, M., LIU, Y., COARFA, C., HARRIS, R. A., SHORESH, N., EPSTEIN, C. B., GJONESKA, E., LEUNG, D., XIE, W., HAWKINS, R. D., LISTER, R., HONG, C., GASCARD, P., MUNGALL, A. J., MOORE, R., CHUAH, E., TAM, A., CANFIELD, T. K., HANSEN, R. S., KAUL, R., SABO, P. J., BANSAL, M. S., CARLES, A., DIXON, J. R., FARH, K. H., FEIZI, S., KARLIC, R., KIM, A. R., KULKARNI, A., LI, D., LOWDON, R., ELLIOTT, G., MERCER, T. R., NEPH, S. J., ONUCHIC, V., POLAK, P., RAJAGOPAL, N., RAY, P., SALLARI, R. C., SIEBENTHALL, K. T., SINNOTT-ARMSTRONG, N. A., STEVENS, M., THURMAN, R. E., WU, J., ZHANG, B., ZHOU, X., BEAUDET, A. E., BOYER, L. A., DE JAGER, P. L., FARNHAM, P. J., FISHER, S. J., HAUSSLER, D., JONES, S. J., LI, W., MARRA, M. A., MCMANUS, M. T., SUNYAEV, S., THOMSON, J. A., TLSTY, T. D., TSAI, L. H., WANG, W., WATERLAND, R. A., ZHANG, M. Q., CHADWICK, L. H., BERNSTEIN, B. E., COSTELLO, J. F., ECKER, J. R., HIRST, M., MEISSNER, A., MILOSAVLJEVIC, A., REN, B., STAMATOYANNOPOULOS, J. A., WANG, T. & KELLIS, M. 2015. Integrative analysis of 111 reference human epigenomes. *Nature,* 518**,** 317-30.

RODRIGUEZ-CAMPOS, A. & AZORIN, F. 2007. RNA is an integral component of chromatin that contributes to its structural organization. *Plos One,* 2**,** e1182.

ROSE, N. R. & KLOSE, R. J. 2014. Understanding the relationship between DNA methylation and histone lysine methylation. *Biochim Biophys Acta,* 1839**,** 1362-72.

SALZ, T., DENG, C., PAMPO, C., SIEMANN, D., QIU, Y., BROWN, K. & HUANG, S. 2014a. Histone methyltransferase hSETD1A is a Novel Regulator of Metastasis in Breast Cancer. *Mol Cancer Res*.

SALZ, T., LI, G., KAYE, F., ZHOU, L., QIU, Y. & HUANG, S. 2014b. hSETD1A regulates Wnt target genes and controls tumor growth of colorectal cancer cells. *Cancer Res,* 74**,** 775-86.

SAPRA, A. K., ANKO, M. L., GRISHINA, I., LORENZ, M., PABIS, M., POSER, I., ROLLINS, J., WEILAND, E. M. & NEUGEBAUER, K. M. 2009. SR protein family members display diverse activities in the formation of nascent and mature mRNPs in vivo. *Mol Cell,* 34**,** 179-90.

SAYEGH, J., CAO, J., ZOU, M. R., MORALES, A., BLAIR, L. P., NORCIA, M., HOYER, D., TACKETT, A. J., MERKEL, J. S. & YAN, Q. 2013. Identification of small molecule inhibitors of Jumonji AT-rich interactive domain 1B (JARID1B) histone demethylase by a sensitive high throughput screen. *J Biol Chem,* 288**,** 9408-17.

SCHRAM, A. W., BAAS, R., JANSEN, P. W., RISS, A., TORA, L., VERMEULEN, M. & TIMMERS, H. T. 2013. A dual role for SAGA-associated factor 29 (SGF29) in ER stress survival by coordination of both histone H3 acetylation and histone H3 lysine-4 trimethylation. *PLoS One,* 8**,** e70035.

SCHWARTZ, S., MESHORER, E. & AST, G. 2009. Chromatin organization marks exon-intron structure. *Nat Struct Mol Biol,* 16**,** 990-5.

SCOTT, L. M. & REBEL, V. I. 2013. Acquired mutations that affect pre-mRNA splicing in hematologic malignancies and solid tumors. *J Natl Cancer Inst,* 105**,** 1540-9.

SHI, X., HONG, T., WALTER, K. L., EWALT, M., MICHISHITA, E., HUNG, T., CARNEY, D., PENA, P., LAN, F., KAADIGE, M. R., LACOSTE, N., CAYROU, C., DAVRAZOU, F., SAHA, A., CAIRNS, B. R., AYER, D. E., KUTATELADZE, T. G., SHI, Y., COTE, J., CHUA, K. F. & GOZANI, O. 2006. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature,* 442**,** 96-9.

SHILATIFARD, A. 2012. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem,* 81**,** 65-95.

SHINSKY, S. A., MONTEITH, K. E., VIGGIANO, S. & COSGROVE, M. S. 2015. Biochemical Reconstitution and Phylogenetic Comparison of Human SET1 Family Core Complexes Involved in Histone Methylation. *J Biol Chem*.

SIMS, R. J., 3RD, MILLHOUSE, S., CHEN, C. F., LEWIS, B. A., ERDJUMENT-BROMAGE, H., TEMPST, P., MANLEY, J. L. & REINBERG, D. 2007. Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment

of transcription postinitiation factors and pre-mRNA splicing. *Mol Cell,* 28**,** 665-76.
SJOLINDER, M., BJORK, P., SODERBERG, E., SABRI, N., FARRANTS, A. K. & VISA, N. 2005. The growing pre-mRNA recruits actin and chromatin-modifying factors to transcriptionally active genes. *Genes Dev,* 19**,** 1871-84.
SPERLING, J., AZUBEL, M. & SPERLING, R. 2008. Structure and function of the Pre-mRNA splicing machine. *Structure,* 16**,** 1605-15.
TAHILIANI, M., MEI, P., FANG, R., LEONOR, T., RUTENBERG, M., SHIMIZU, F., LI, J., RAO, A. & SHI, Y. 2007. The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. *Nature,* 447**,** 601-5.
TATE, C. M., LEE, J. H. & SKALNIK, D. G. 2010. CXXC finger protein 1 restricts the Setd1A histone H3K4 methyltransferase complex to euchromatin. *FEBS J,* 277**,** 210-23.
TEOH, P. L. & SHARROCKS, A. D. 2014. WDR5, ASH2L, and RBBP5 control the efficiency of FOS transcript processing. *Cell Mol Biol Lett,* 19**,** 215-32.
TERRENOIRE, E., MCRONALD, F., HALSALL, J. A., PAGE, P., ILLINGWORTH, R. S., TAYLOR, A. M., DAVISON, V., O'NEILL, L. P. & TURNER, B. M. 2010. Immunostaining of modified histones defines high-level features of the human metaphase epigenome. *Genome Biol,* 11**,** R110.
THINNES, C. C., ENGLAND, K. S., KAWAMURA, A., CHOWDHURY, R., SCHOFIELD, C. J. & HOPKINSON, R. J. 2014. Targeting histone lysine demethylases - progress, challenges, and the future. *Biochim Biophys Acta,* 1839**,** 1416-32.
THOMSON, J. P., SKENE, P. J., SELFRIDGE, J., CLOUAIRE, T., GUY, J., WEBB, S., KERR, A. R., DEATON, A., ANDREWS, R., JAMES, K. D., TURNER, D. J., ILLINGWORTH, R. & BIRD, A. 2010. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature,* 464**,** 1082-6.
TILGNER, H., NIKOLAOU, C., ALTHAMMER, S., SAMMETH, M., BEATO, M., VALCARCEL, J. & GUIGO, R. 2009. Nucleosome positioning as a determinant of exon recognition. *Nat Struct Mol Biol,* 16**,** 996-1001.
TSAI, M. C., MANOR, O., WAN, Y., MOSAMMAPARAST, N., WANG, J. K., LAN, F., SHI, Y., SEGAL, E. & CHANG, H. Y. 2010. Long noncoding RNA as modular scaffold of histone modification complexes. *Science,* 329**,** 689-93.
TSAI, W. W., NIESSEN, S., GOEBEL, N., YATES, J. R., 3RD, GUCCIONE, E. & MONTMINY, M. 2013. PRMT5 modulates the metabolic response to fasting signals. *Proc Natl Acad Sci U S A,* 110**,** 8870-5.
TU, S., TENG, Y. C., YUAN, C., WU, Y. T., CHAN, M. Y., CHENG, A. N., LIN, P. H., JUAN, L. J. & TSAI, M. D. 2008. The ARID domain of the H3K4 demethylase RBP2 binds to a DNA CCGCCC motif. *Nat Struct Mol Biol,* 15**,** 419-21.
TURNER, B. M. 2014. Nucleosome signalling; an evolving concept. *Biochim Biophys Acta,* 1839**,** 623-6.

Draft

TYAGI, A., RYME, J., BRODIN, D., OSTLUND FARRANTS, A. K. & VISA, N. 2009. SWI/SNF associates with nascent pre-mRNPs and regulates alternative pre-mRNA processing. *PLoS Genet,* 5**,** e1000470.
TZSCHACH, A., LENZNER, S., MOSER, B., REINHARDT, R., CHELLY, J., FRYNS, J. P., KLEEFSTRA, T., RAYNAUD, M., TURNER, G., ROPERS, H. H., KUSS, A. & JENSEN, L. R. 2006. Novel JARID1C/SMCX mutations in patients with X-linked mental retardation. *Hum Mutat,* 27**,** 389.
VAN NULAND, R., SMITS, A. H., PALLAKI, P., JANSEN, P. W., VERMEULEN, M. & TIMMERS, H. T. 2013. Quantitative dissection and stoichiometry determination of the human SET1/MLL histone methyltransferase complexes. *Mol Cell Biol,* 33**,** 2067-77.
VAN RECHEM, C. & WHETSTINE, J. R. 2014. Examining the impact of gene variants on histone lysine methylation. *Biochim Biophys Acta*.
VETHANTHAM, V., YANG, Y., BOWMAN, C., ASP, P., LEE, J. H., SKALNIK, D. G. & DYNLACHT, B. D. 2012. Dynamic loss of H2B ubiquitylation without corresponding changes in H3K4 trimethylation during myogenic differentiation. *Mol Cell Biol,* 32**,** 1044-55.
WAHL, M. C., WILL, C. L. & LUHRMANN, R. 2009. The spliceosome: design principles of a dynamic RNP machine. *Cell,* 136**,** 701-18.
WANG, Z., ZANG, C., CUI, K., SCHONES, D. E., BARSKI, A., PENG, W. & ZHAO, K. 2009. Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell,* 138**,** 1019-31.
WEIRICH, S., KUDITHIPUDI, S., KYCIA, I. & JELTSCH, A. 2015. Somatic cancer mutations in the MLL3-SET domain alter the catalytic properties of the enzyme. *Clin Epigenetics,* 7**,** 36.
WIJEWEERA, A., HAJ, M., FELDMAN, A., PNUELI, L., LUO, Z. & MELAMED, P. 2015. Gonadotropin gene transcription is activated by menin-mediated effects on the chromatin. *Biochim Biophys Acta*.
WONG, P. P., MIRANDA, F., CHAN, K. V., BERLATO, C., HURST, H. C. & SCIBETTA, A. G. 2012. Histone demethylase KDM5B collaborates with TFAP2C and Myc to repress the cell cycle inhibitor p21(cip) (CDKN1A). *Mol Cell Biol,* 32**,** 1633-44.
WU, L., LEE, S. Y., ZHOU, B., NGUYEN, U. T., MUIR, T. W., TAN, S. & DOU, Y. 2013. ASH2L regulates ubiquitylation signaling to MLL: trans-regulation of H3 K4 methylation in higher eukaryotes. *Mol Cell,* 49**,** 1108-20.
XIANG, Y., ZHU, Z., HAN, G., YE, X., XU, B., PENG, Z., MA, Y., YU, Y., LIN, H., CHEN, A. P. & CHEN, C. D. 2007. JARID1B is a histone H3 lysine 4 demethylase up-regulated in prostate cancer. *Proc Natl Acad Sci U S A,* 104**,** 19226-31.
XIAO, N., LAHA, S., DAS, S. P., MORLOCK, K., JESNECK, J. L. & RAFFEL, G. D. 2014. Ott1(Rbm15) regulates thrombopoietin response in hematopoietic stem cells through alternative splicing of c-Mpl. *Blood*.
XU, C., BIAN, C., LAM, R., DONG, A. & MIN, J. 2011. The structural basis for selective binding of non-methylated CpG islands by the CFP1 CXXC domain. *Nat Commun,* 2**,** 227.

Draft

YAMANE, K., TATEISHI, K., KLOSE, R. J., FANG, J., FABRIZIO, L. A., ERDJUMENT-BROMAGE, H., TAYLOR-PAPADIMITRIOU, J., TEMPST, P. & ZHANG, Y. 2007. PLU-1 is an H3K4 demethylase involved in transcriptional repression and breast cancer cell proliferation. *Mol Cell,* 25**,** 801-12.

YOKOYAMA, A. 2015. Molecular mechanisms of MLL-associated leukemia. *Int J Hematol,* 101**,** 352-61.

YUAN, C. C., MATTHEWS, A. G., JIN, Y., CHEN, C. F., CHAPMAN, B. A., OHSUMI, T. K., GLASS, K. C., KUTATELADZE, T. G., BOROWSKY, M. L., STRUHL, K. & OETTINGER, M. A. 2012. Histone H3R2 symmetric dimethylation and histone H3K4 trimethylation are tightly correlated in eukaryotic genomes. *Cell Rep,* 1**,** 83-90.

ZAIDI, S. K., GRANDY, R. A., LOPEZ-CAMACHO, C., MONTECINO, M., VAN WIJNEN, A. J., LIAN, J. B., STEIN, J. L. & STEIN, G. S. 2014. Bookmarking target genes in mitosis: a shared epigenetic trait of phenotypic transcription factors and oncogenes? *Cancer Res,* 74**,** 420-5.

ZAIDI, S. K., YOUNG, D. W., MONTECINO, M., VAN WIJNEN, A. J., STEIN, J. L., LIAN, J. B. & STEIN, G. S. 2011. Bookmarking the genome: maintenance of epigenetic information. *J Biol Chem,* 286**,** 18355-61.

ZHOU, H. L., LUO, G., WISE, J. A. & LOU, H. 2014. Regulation of alternative splicing by local histone modifications: potential roles for RNA-guided mechanisms. *Nucleic Acids Res,* 42**,** 701-13.

Draft

Figure Legends

Figure 1. *MCL1* has two peaks of H3K4me3. One peak is at the 5' splice site of the first exon, which has a CpG island, and a second peak is over exon 2. SETD1A &/or B are recruited via CXXC1 (CFP1) subunit to CpG islands. The H3K4me3 modified nucleosomes bind to multiple readers (CHD1, U2AF65, U2 snRNP, PAF, FACT and KATs). CHD1, KAT2B, U2 snRNP, U2AF65 and HDAC1/2 are bound to the transcript and from the RNA platform dialogue with the H3K4me3 modified nucleosomes. KDM5 family members are associated with HDAC1 & 2. The SETD1A complex is from (van Nuland et al., 2013) and the pre-mRNA splicing events are from (Braunschweig et al., 2013). The definition of a CpG island is a region with at least 200 bp and GC percentage that is greater than 50%, and with an observed to expected CpG ratio that is greater than 60%. M, methylation; Ac, acetylation.

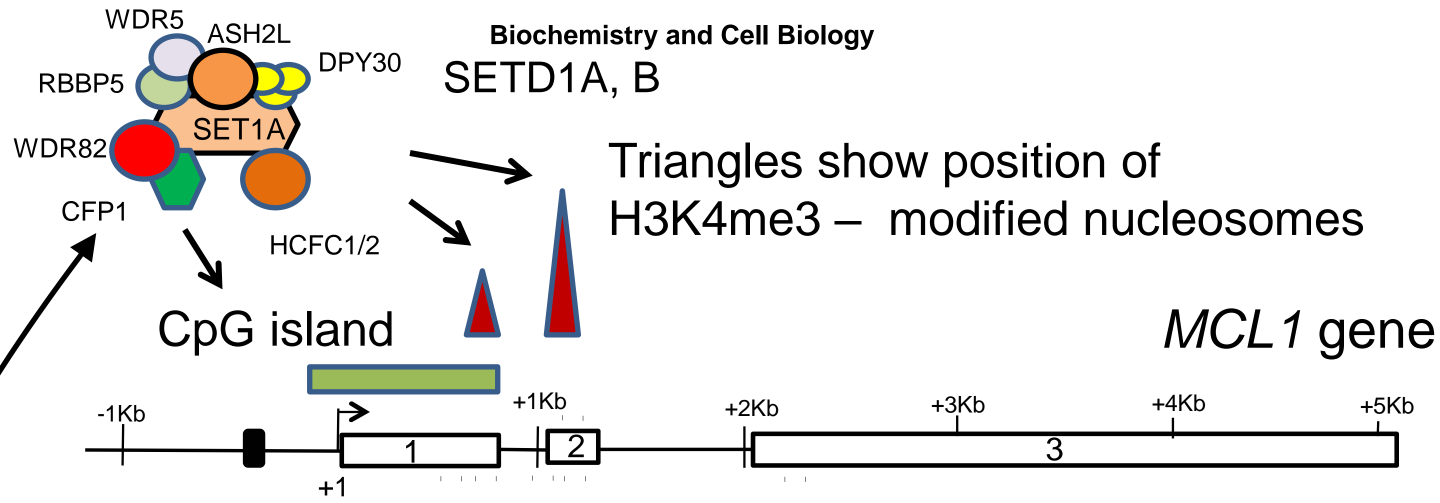
Draft

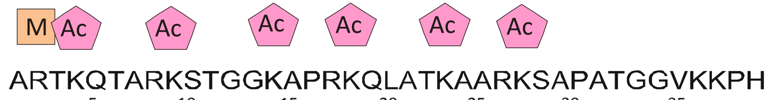
Figure 2. Location of H3K4me3 relative to CpG islands and the 5' splice site of the first exon of *RBM23*, *PRMT5*, *FOS*, *DUSP1* and *MCL1* genes. Direction of transcription is indicated by arrow. Tracks are from Encode. http://genome.ucsc.edu/ENCODE/dataMatrix/encodeDataMatrixHuman.html). Note that H3K4me3 is found over the promoter region but the stronger H3K4me3 peaks are found in the gene body of transcribed genes in HCT116 cells. The tracks for the *RBM23* and *PRMT5* genes in HCT116 and A549 cells are shown.

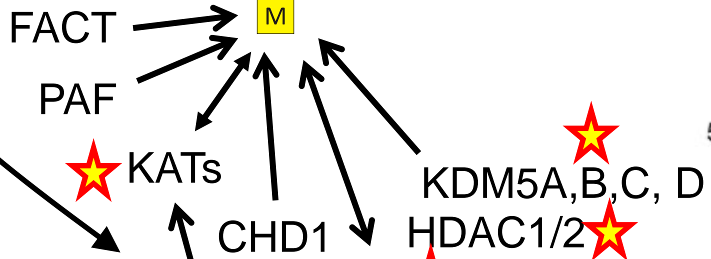
Figure 3. Location of H3K4me3 in the *APOL2*, *APOL1*, and *TFF3* genes, which lack CpG islands, in HepG2, HCT116 and A549 cells. Direction of transcription is indicated by arrow. Tracks are from Encode.

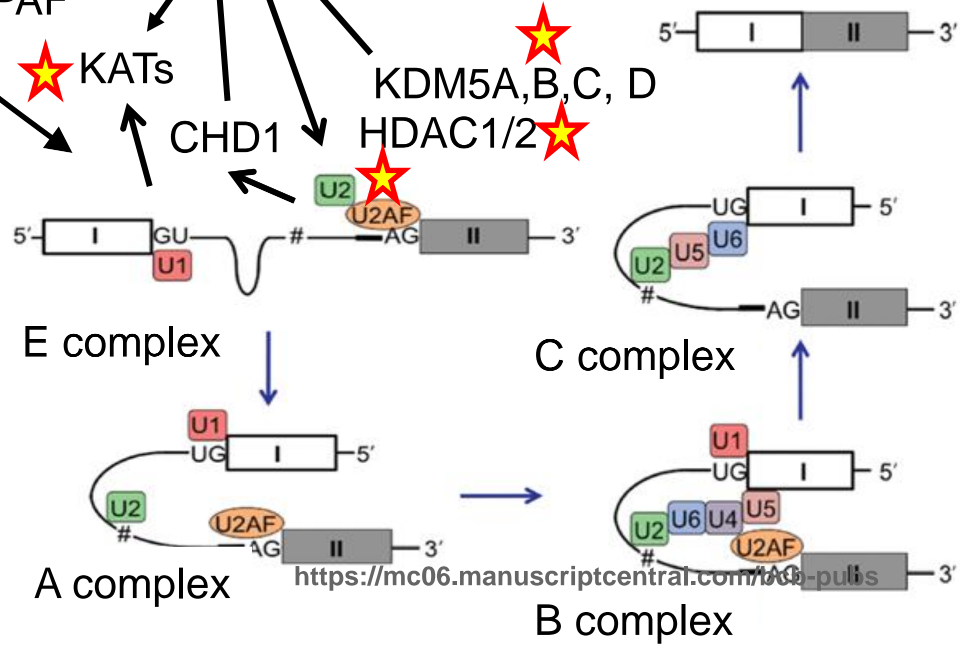
Figure 4. Location of H3K4me3 along the *MPL* gene body in HCT116, A549, MCF7 and HeLa cells. For HeLa cells the location of H3K36me3 is also shown. Direction of transcription is indicated by arrow. Tracks are from Encode.

Draft





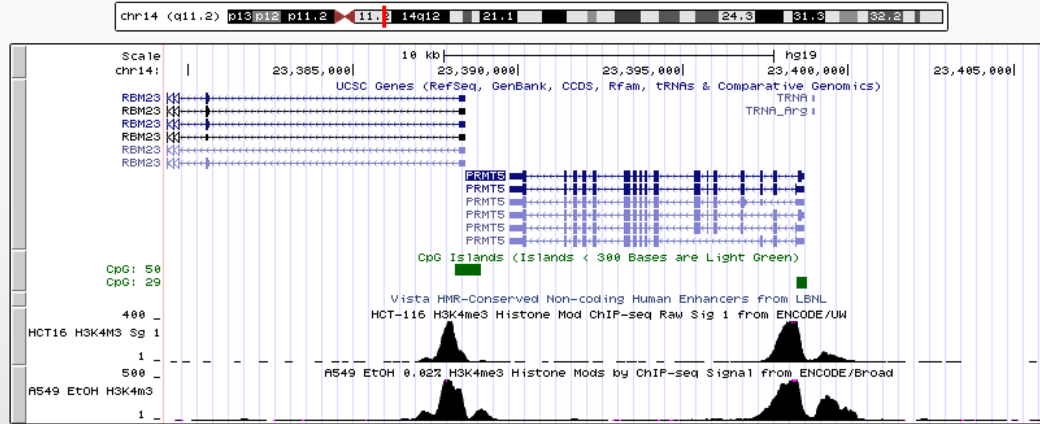


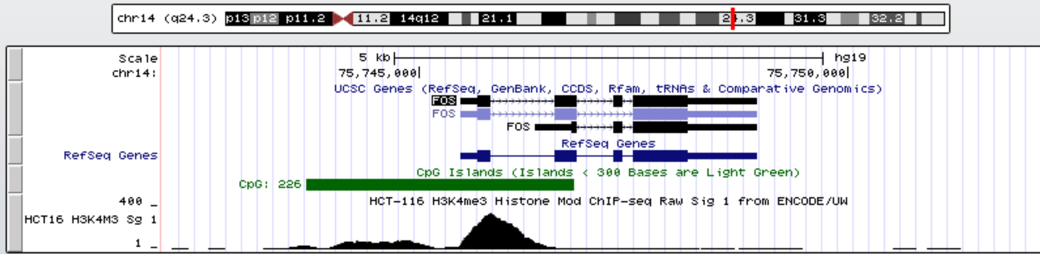


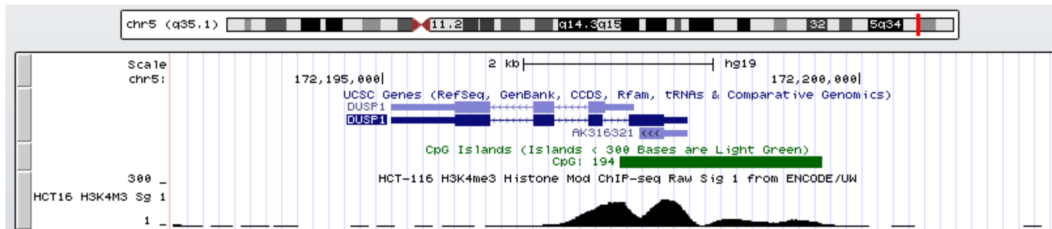
Over-expressed or mutated in breast cancer

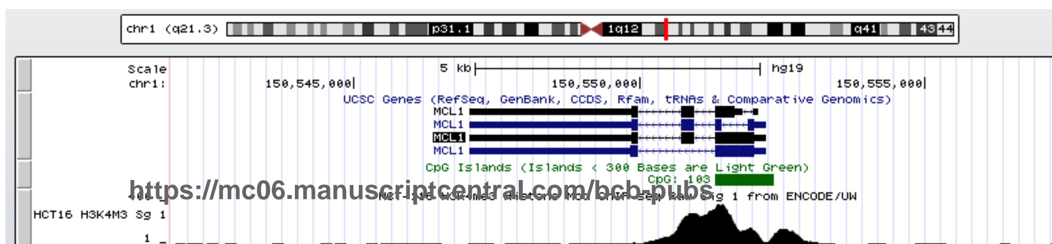
https://mc06.manuscriptcentral.com/bcb-pubs

Fig. 2



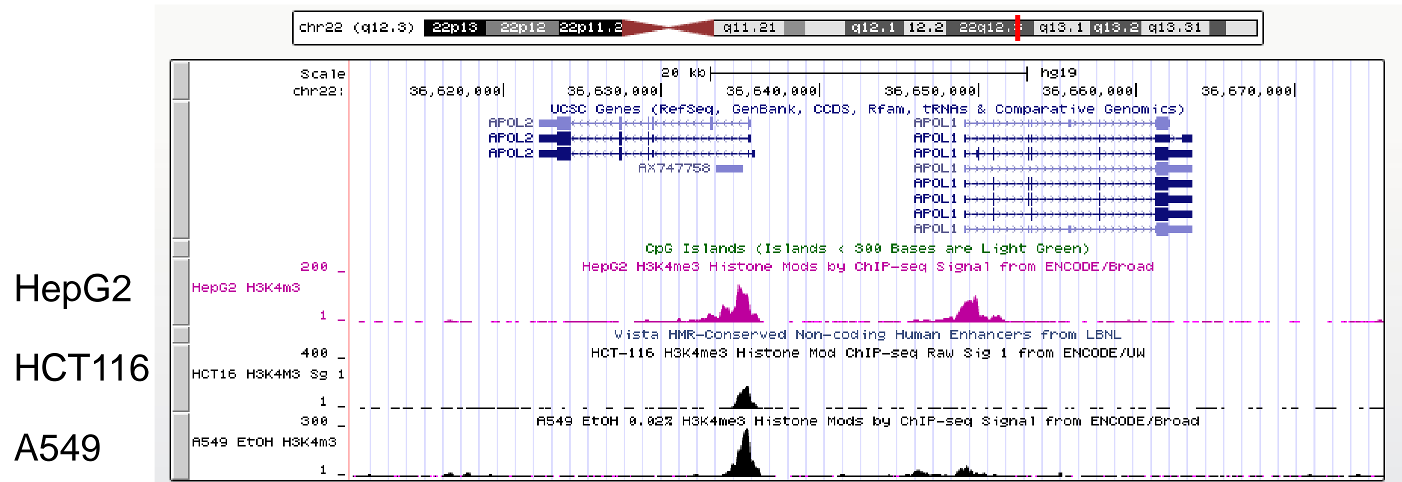






https://mc06.manuscriptcentral.com/bcb-pubs

Fig. 3



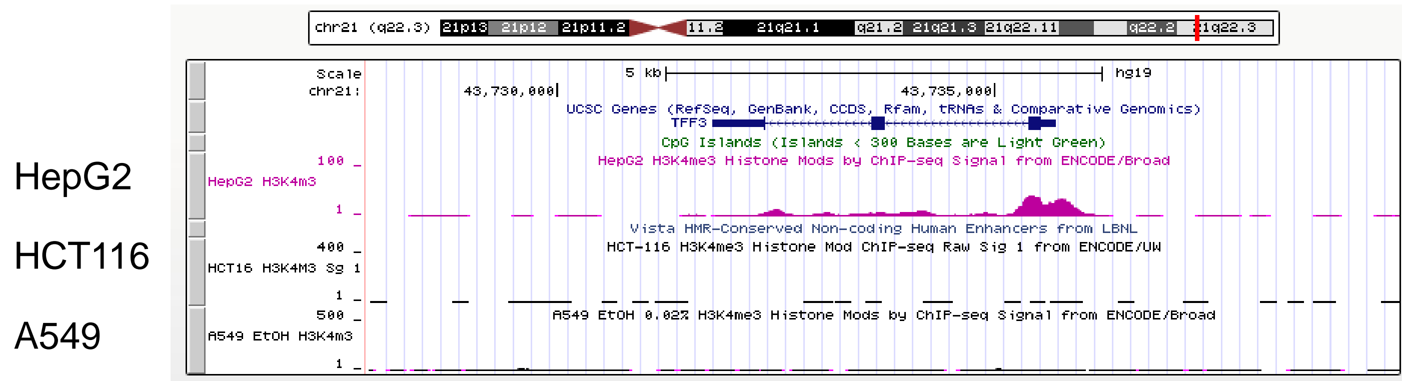


Fig. 4

MPL, MPL proto-oncogene, thrombopoietin receptor →

HCT116

A549

MCF7

HeLa

HeLa