Published in final edited form as:

Expert Rev Proteomics. 2019 January ; 16(1): 33–47. doi:10.1080/14789450.2019.1550363.

Proteomic approaches for cancer epigenetics research

Dylan M. Marchione¹, Benjamin A. Garcia¹, and John Wojcik²

¹Epigenetics Institute, Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States.

²Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States.

Abstract

Introduction: Epigenetic dysregulation drives or supports numerous human cancers. The chromatin landscape in cancer cells is often marked by abnormal histone modification patterns (PTMs) and by aberrant assembly and recruitment of protein complexes to specific genomic loci. Mass spectrometry-based proteomic analyses can support the discovery and characterization of both phenomena.

Areas covered: We broadly divide this literature into two parts: “modification-centric” analyses that link histone PTMs to cancer biology, and “complex-centric” analyses, that examine protein-protein interactions that occur *de novo* as a result of oncogenic mutations. We also discuss proteomic studies of oncohistones. We highlight relevant examples, discuss limitations, and speculate about forthcoming innovations regarding each application.

Expert commentary: “Modification-centric” analyses have been used to further understanding of cancer’s histone code and to identify associated therapeutic vulnerabilities. “Complex-centric” analyses have likewise revealed insights into mechanisms of oncogenesis and suggested potential therapeutic targets, particularly in MLL-associated leukemias. Proteomic experiments have also supported some of the pioneering studies of oncohistone-mediated tumorigenesis. Additional applications of proteomics that may benefit cancer epigenetics research include middle-down and top-down histone PTM analysis, chromatin reader profiling, and genomic locus-specific protein identification. In the coming years, proteomic approaches will remain powerful ways to interrogate the biology of cancer.

Graphical abstract

Correspondence: John Wojcik, Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States, John.Wojcik@uphs.upenn.edu.

Declaration of interest

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

Reviewer disclosures

Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

Keywords

Affinity proteomics; Cancer; Chromatin; Epigenetics; Histones; Readers; Middle-down; Top-down

1. Introduction

Basic research on cancer often begins with the discovery of a recurrent disease-associated mutation and proceeds with a functional characterization of the mutant gene product or products. Mutations are then contextualized based on how they change the wild-type gene product: they may result in copy number variation, loss- or gain-of-function, or they may be more complex. Classically, they are then classified as either proto-oncogenes or tumor suppressors based on whether they directly promote growth or instead release cells from growth restriction.

Collectively, epigenetic regulators are among the most frequently mutated genes in cancer [1]. However, they do not fit neatly into this paradigm. A recent report suggested that many should be classified as “proto-oncogenes with tumor suppressor function,” indicating that their relationship to oncogenesis is more complex [2]. Moreover, their expression is often dysregulated in cancer in the absence of mutations, leading Vogelstein and Kinzler to ask whether they are “major cast members” or “bit players” [3]. They appear to be the former, occurring both as a cause and a consequence of tumor evolution [4,5]. In general, the mechanisms by which these genes contribute to cancer are typically subtler than those of classic oncogenes and tumor suppressors.

Rather than altering signal transduction or cell cycle progression directly, the epigenetic regulators we discuss here are those that modulate transcription by altering chromatin accessibility or associations, thereby influencing the expression of many genes across diverse pathways. This is typically achieved through alterations in the positioning or post-translational modifications (PTMs) of histones, the small, basic proteins that contribute to DNA condensation and supercoiling. The N-terminal tails of histones, especially of H3 and H4, are rich in amino acid residues that can be modified by a wide range of PTMs including mono-, di-, and trimethylation (me1, me2, me3), acetylation (ac), phosphorylation (ph), ubiquitination (ub), and at least 10 other functional groups [6].

The histone code hypothesis is the viewpoint that collectively, the pattern of these modifications around a particular genetic locus represents a set of instructions that are interpretable by chromatin-interacting (“reader”) proteins, and perhaps one day by human investigators [7–9]. Indeed, certain histone modifications have been clearly associated with specific functions: histone H3K4me3 typically marks promoters of actively transcribed genes [10], H3K9me3 marks constitutive heterochromatin [11], H3K27me3 marks facultative heterochromatin [12], and H3K27ac marks active enhancers [13]. Current approaches to decipher the histone code are not unlike those which are used to decipher any other unknown language: the strategy favored by descriptive linguists is to (1) define the lexicon and (2) characterize the grammar. In this case, the lexicon is the entire catalogue of histone modifications that can be utilized by cells. The grammar, then, is the principle of organization that dictates when specific modifications are used and how they relate to each other. MS is a uniquely well-suited approach to achieve both of these aims: targeted MS analyses have been used to discover novel histone PTMs such as crotonylation [14], 2-hydroxyisobutyrylation [15], and benzoylation [16], and middle-down and top-down proteomics analyses have been used to systematically catalogue recurrent PTM combinations [17–20].

Briefly, we define “proteomics” as any MS-based effort to comprehensively characterize the proteins and/or protein post-translational modifications (PTMs) in a sample. In a traditional bottom-up proteomics experiment, proteins are digested with one or more proteases in order to generate short (~6–18 amino acid) peptides that are amenable to chromatographic separation. Middle-down proteomics refers to the approach in which proteins are digested into longer (~50 amino acid) peptides. Middle-down offers a few specific advantages over bottom-up proteomics: because it entails the analysis of longer peptides, it results in greater protein sequence coverage, often allowing for evaluation of combinatorial PTMs and detection of protein isoforms [17,21]. Middle-down has not been widely adopted yet, likely due in part to the paucity of computational software available for interpreting the complex spectra that result [22]. However, it is steadily gaining traction – especially in the context of histone modification analysis [18,23,24]. In top-down strategies, intact proteins are infused into a mass spectrometer. The main strength of top-down proteomics is that by definition it guarantees 100% sequence coverage. It is the only MS-based strategy capable of analyzing “proteoforms” directly, regardless of the presence of genetic polymorphisms, alternative splicing, PTMs, or any other expected or unexpected alteration. Like middle-down, top-down proteomics remains technically challenging, but it is constantly becoming more accessible [25].

In this review, we aim to convey how the general proteomics workflow has been adapted to decipher epigenetic mechanisms in cancer. One specialized application of proteomics which has contributed to the understanding of epigenetic mechanisms associated with various cancers is the specific enrichment of histones for PTM analysis. Another variant is affinity proteomics (AP-MS or IP-MS), wherein a protein of interest is immunoprecipitated from a lysate in order to identify any interacting proteins [26]. This approach has been essential in deciphering the oncogenic mechanisms associated with a subset of MLL translocation driven leukemias.

Please note, the focus of this review is not on general proteomics (see [27,28]), on the proteomic analysis of histones (see [29–31]) nor on epigenetic mechanisms of cancer (see [1,32,33]). This review examines the specific application of proteomics to cancer epigenetics research. We review the history, highlight important lessons learned via several “case reports,” discuss outstanding issues, and conclude with an outlook regarding the future of this exciting and fruitful area of research.

2. Modification-centric analysis

2.1. A common histone code for cancer?

In 1950, Edgar and Ellen Stedman reported that histones isolated from different tissue types had different biochemical properties. They proposed that “cell-specific” histones might mediate the selective gene expression necessary for functional specialization [34]. Before long, H.J. Cruft and colleagues reasoned that if this hypothesis were correct, then “the transformation of a normal into a malignant cell would necessarily be associated with a change in the histone which it contained” [35]. Indeed, they reported that histones isolated from six distinct cancer cell types had a markedly lower isoelectric point than histones isolated from normal cells. They concluded that “there is presumably an underlying chemical cause of these differences, but its nature has not yet been ascertained” [35].

Just 10 years later, it was determined that histones can be post-translationally modified by acetylation, and that acetylated histones were more permissive of transcription than unmodified histones [36]. The following year, Jay Roth hypothesized that cancer might be caused by “interference...with the enzyme system or systems responsible for acetylation or other chemical modification of histones” [37]. He speculated that dysregulation of histone modifications might drive tumorigenesis by promoting aberrant expression of some genes, and repression of others, thereby promoting growth and inhibiting differentiation. It would be nearly 40 years until this remarkably perspicacious hypothesis was again articulated and examined.

In the intervening years, advances in LC-MS made it possible to study histone modifications comprehensively, providing the means necessary to resuscitate Roth’s dormant hypothesis. In 2004, Mario Fraga and colleagues performed top-down LC-MS and high-performance capillary electrophoresis to analyze the PTMs present on isolated histone H4 in a large panel of cancer cell lines and primary tumors, as well as an assortment of controls, with the specific goal of identifying abnormal PTMs in the cancer cells. Histone H4 isolated from cancer cells displayed a striking reduction in both monoacetylation and trimethylation. These changes were also identified in a mouse model of skin cancer progression and their magnitude correlated with the cancer stage. Using tandem MS (MS/MS) and western blotting, they determined that the loss of acetylation primarily occurred on lysine 16 (H4K16ac) and the loss of trimethylation occurred on lysine 20 (H4K20me3). Interestingly, despite the loss of H4K16ac, the cancer cells generally had an increase in triacetylated H4, perhaps explaining the lower isoelectric point first observed by Cruft et al. While the relationship between these PTM changes and oncogenesis was not clearly established, this study was an important conceptual first step in the use of MS to profile the global chromatin landscape of cancer cells [38].

Several subsequent MS-based studies also aimed to identify shared epigenetic changes across cancer cells. One study used stable isotope labeling and LC-MS/MS to comprehensively analyze the modifications on both histones H3 and H4 from a variety of breast cancer lines. Their results similarly demonstrated reduced levels of H4K16ac and H4K20me3 in the cancer cells relative to normal breast epithelium. They also revealed other consistent epigenetic changes, most notably an increase in H3K27me2/3 [39]. A subsequent analysis of 24 different cell lines likewise demonstrated a consistent elevation in H3K27me2/3 in cancer cells [40]. Interestingly, the latter study also included a microarray analysis of 224 histone-modifying genes, and found that the abundance of a given histone PTM did not always correlate with the abundance of the transcripts of the enzymes that regulate it, highlighting the importance of measuring the PTMs directly. The study also reported that two breast cancer lines displaying high levels of the *EZH2* transcript and its catalytic byproducts H3K27me2/3 were particularly sensitive to *EZH2* knockdown, demonstrating that profiling either the expression of histone modifiers, the abundance of specific histone PTMs, or both, might predict therapeutic vulnerabilities. While the underlying mechanism was not established, one way to do so would be to map where in the genome the PTM changes were localized via chromatin immunoprecipitation and high throughput sequencing (ChIP-seq), and then to cross-reference that data with a transcriptome analysis to identify affected genes [40]. Indeed, this has proven to be a powerful approach. An example of the suggested workflow is depicted in Figure 1.

The merit of these earlier studies notwithstanding, it is highly unlikely that there is a singular cancer-associated histone PTM signature, just as there is no such thing as a singular cancer-associated transcriptome. The observation that there are certain histone-modifying enzymes or complexes for which both oncogenic loss-of-function and gain-of-function mutations have been observed – as is the case with *EZH2* – is strong evidence against this hypothesis [2]. Instead, the histone PTM landscape in a given cancer cell type is likely to be the result of a complex interplay between its cell-of-origin, its expression of histone modifying enzymes, its mutational landscape, and its metabolism. Nonetheless, the systematic measurement of histone PTMs in a particular cancer type can often suggest modifications that might be related to oncogenic gene expression programs. Because the enzymes responsible for most PTMs are known and increasingly becoming druggable, this approach can implicate candidate phenotypic drivers, and occasionally, avenues for therapeutic intervention (reviewed in [41]). A selection of histone PTMs for which inhibitors are available is included in Table 1.

Importantly, a recent study that compared the histone PTM profiles of cancer cell lines, primary cells grown in culture, and formalin-fixed paraffin-embedded (FFPE) tumor tissue revealed that there are marked and time-dependent epigenetic changes that occur in culture, affecting both methylation and acetylation. Interestingly, these changes largely reverted when cancer cell lines were implanted as xenografts [53]. These results strongly suggest that at a minimum, epigenetic findings in cell culture models should be recapitulated in xenograft models, but the ideal approach would be to utilize FFPE tissue. To this end, a method to extract histones from FFPE for MS analysis was recently described, and has since been adapted to utilize laser microdissection to isolate specific cell populations [54,55]. Despite

its limitations, however, MS-based histone PTM profiling in cancer cell lines has shed light on several oncogenic mechanisms.

2.2. NSD2/MMSET/WHSC1

Multiple myeloma (MM) provides an illustrative example in which histone PTM analysis helped determine the oncogenic mechanism of an epigenetically-driven cancer. Approximately 15–20% of MM cases possess a t(4;14) chromosomal translocation that results in an aberrant upregulation of two genes: the histone lysine methyltransferase *NSD2/MMSET/WHSC1*, hereafter referred to as *NSD2*, and the fibroblast growth factor receptor *FGFR3*. In the initial report, both gene products were hypothesized to have “oncogenic potential” but the relative contribution of each was unknown [56]. Later, others found that the chromosomal translocation was an adverse prognostic factor in MM irrespective of *FGFR3* expression, suggesting that *NSD2* might be the main driver of oncogenesis [57]. The mechanism for this, however, was uncertain, because the enzymatic function of NSD2 was ambiguous: it has been separately reported as a methyltransferase for histone H3K4, H3K27, H3K36, and H4K20 [58–63].

A careful MS-based study elucidated the mechanism underlying translocation-associated MM by first clarifying that the principal intracellular byproduct of NSD2 was H3K36me2. Moreover, the authors showed that expression of a single copy of the fusion gene was sufficient to increase cellular H3K36me2 approximately 3-fold. ChIP-seq revealed that the fusion gene also significantly disrupted the genomic distribution of H3K36me2. Correlation of the ChIP-seq data with microarray gene expression data demonstrated that the genes with the greatest gain of H3K36me2 were also the most upregulated, and this gene set was significantly enriched for genes in numerous cancer pathways. Accordingly, expression of a constitutively active form of NSD2 was sufficient to promote increased proliferation, anchorage-independent growth, and xenograft tumor formation in t(4;14) negative myeloma cells. Thus, the proteomic analysis of histone PTM patterns was a critical first step towards determining the oncogenic mechanism of t(4;14) translocation driven MM [64].

2.3. SIRT7

SIRT7 is a sirtuin, a member of a family of proteins named for their homology to the yeast NAD⁺-dependent histone deacetylase Sir2 (silent information regulator-2). In 2005, it was reported that SIRT7 is primarily localized to heterochromatin and nucleoli, suggesting that it might contribute to epigenetic regulation [65]. In 2012, an *in vitro* screen of recombinant SIRT7 against 12 different acetylated histone peptides followed by LC-MS/MS revealed that SIRT7 robustly and specifically deacetylates H3K18ac. This was also confirmed with purified poly-nucleosomes [66]. H3K18ac was previously known to occupy promoter regions and correlate with transcriptional activation [67]. Interestingly, low levels of H3K18ac had been associated with poor outcomes in lung, kidney, and pancreatic cancers [68,69]. Together these observations suggested that SIRT7-mediated H3K18 deacetylation might contribute to the repression of a subset of genes in cancer. ChIP-sequencing of SIRT7 in HT1080 (fibrosarcoma) cells identified 276 binding sites, many of which were previously known to be modified by H3K18ac. Among these sites, 241 corresponded to protein-coding genes, and this gene set was highly enriched for proteins with tumor suppressive function.

Subsequent work demonstrated that deacetylation of H3K18 by SIRT7 was essential for anchorage-independent growth and resistance to contact inhibition in HT1080 and U2OS (osteosarcoma) cells *in vitro*, and for tumor formation in a mouse xenograft model using U251 (glioma) cells [66]. Collectively the results suggested that SIRT7-mediated deacetylation of H3K18 may represent a common mechanism exploited by cancer cells to silence tumor suppressors.

2.4. SMYD3

SMYD3 (SET- and MYND-domain containing 3) is a lysine methyltransferase that is highly expressed in a variety of cancers including colorectal, liver, and breast [70–72]. SMYD3 overexpression in NIH3T3 (murine fibroblast) cells drives increased proliferation and colony formation in soft agar [70]. Conversely, SMYD3 knockdown inhibits the proliferation of various cancer cell lines *in vitro* [70,71]. The initial characterization of SMYD3 suggested that it was a H3K4 methyltransferase [70] but a subsequent study argued that its primary substrate was instead H4K20 [73].

To definitively characterize the substrate preference of SMYD3, Van Aller and colleagues performed an *in vitro* screen of recombinant full-length SMYD3 against a library of 327 different histone peptides using ³H-labeled S-adenosyl methionine. They found that SMYD3 robustly methylated H4 peptides spanning positions 1–21, but did not significantly methylate histone H3. After the experiments were replicated with full-length histones and reconstituted nucleosomes as substrates, LC-MS/MS was used to localize the SMYD3-mediated H4 methylation. Unexpectedly, the methylation was primarily localized at H4K5 rather than K20 [72]. It was therefore hypothesized that the function of SMYD3 in cancer was related to H4K5 methylation.

In fact, H4K5 methylation is no longer believed to be the physiologically important substrate of SMYD3 [74]. In 2014, a follow-up study to that of Van Aller reported that SMYD3 is predominantly cytoplasmic in the cancers in which it is most highly expressed, suggesting that it might have non-histone substrates. To identify these substrates, investigators conducted a broader *in vitro* SMYD3 methylation screen utilizing a library of nearly 10,000 proteins. This and follow-up experiments revealed that SMYD3 methylates the MAP-kinase MAP3K2 with a catalytic efficiency nearly two orders of magnitude greater than for H4K5. A series of mechanistic experiments then showed that the ability of SMYD3 to potentiate Ras signaling was largely dependent on its ability to methylate MAP3K2 at K260. These findings fit neatly with the observation that SMYD3 expression is highest in Ras-driven cancers [74].

Interestingly, methylation of MAP3K2 did not alter its kinase activity. In order to address whether methylation instead altered the interaction between MAP3K2 and other proteins, the investigators performed a peptide pulldown proteomics experiment. Briefly, this entailed the design and synthesis of biotinylated peptides corresponding to a 26 amino-acid stretch of MAP3K2 spanning the methylation site. One peptide was unmodified and one was trimethylated at the lysine corresponding to K260. The peptides were then immobilized on streptavidin beads and incubated with HeLa cytoplasmic protein extracts generated through stable isotope labeling by amino acids in cell culture (SILAC) that were either “light”

(unlabeled) or “heavy” ($^2\text{H}_4$ -lysine, $^{13}\text{C}_6$ -arginine-labeled). The two sets of beads were pooled and a traditional shotgun proteomics analysis was performed. In experiments like these, the heavy-to-light intensity ratio of a protein signifies its relative affinity for the two baits. This experiment revealed that the methylation of MAP3K2 did not recruit specific binders, but instead prevented the binding of subunits of the PP2A serine/threonine phosphatase. SMYD3 was therefore hypothesized to act as a MAP3K2 phosphatase inhibitor, potentiating Ras signaling. In support of this hypothesis, treatment with the PP2A inhibitor cantharidin largely phenocopied SMYD3 overexpression [74].

Though it underscores the value of untargeted MS-based proteomic approaches, the progress in understanding the role of SMYD3 in cancer also serves as a warning of the hazards of relying on *in vitro* screens for elucidating enzyme function. Multiple distinct efforts to characterize SMYD3 substrates relied on *in vitro* assays conducted with solely histone substrates [70,72,73]. They were all limited by the assumption that the important substrate for SMYD3 would be a histone, based on the fact that it has SET and MYND domains. The best candidate substrate to date was only identified in a broader screen [74]. Moreover, it remains possible that a complete exploration of the “methylome” of cells in which SMYD3 is inhibited, knocked out, or otherwise inactivated could reveal additional substrates or functions.

2.5. Pitfalls and alternatives related to the modification-centric approaches

As exemplified by the history of SMYD3, one pitfall of histone-centric proteomic analysis is that by design it does not address modifications that exist on nonhistone proteins. This is an important fact to consider given that many enzymes named or annotated as histone modifiers have important non-histone substrates [75,76]. In addition, this also means that many drugs that are named for their ability to inhibit such enzymes might exert important effects through these nonhistone substrates. Indeed, the burgeoning field of chemical proteomics is rapidly reshaping our understanding of such drugs (reviewed in [77]). Therefore, it may be informative to ask whether modifications of nonhistone proteins are changing by profiling the relevant modification proteome-wide.

An additional conceivable pitfall is that there may be cancer-associated mutations that undermine histone PTM crosstalk without driving overt changes in the abundance of any particular modification. Histone PTM crosstalk is, in some cases, genetically encoded. For example, although the genomic distribution of H3K4me3 and H3K27me3 overlap significantly, in regions of “bivalent chromatin,” [78], these marks occupy different copies of H3 in the same nucleosome rather than the same histone tail [79]. This is because methylation of H3K4 inhibits K27 methylation by preventing binding of the Nurf55-Suz12 subcomplex of PRC2 [80]. Similarly, PRC2-mediated H3K27 methylation is significantly inhibited by K36 methylation, unless the PRC2 complex contains the polycomb-like proteins PHF1 or PHF19 [80–82]. It is therefore conceivable that mutations in these methylation-sensing domains of PRC2 subunits, or in histone H3 itself, could disrupt PTM crosstalk mechanisms and, consequently, regulation of gene expression. Because histone tails are digested into many small peptides in bottom-up analysis, this approach would be unlikely

likely to capture disruption of crosstalk that occurs at a distance, but middle-down or top-down strategies could be especially valuable here.

Perhaps a more serious pitfall is the possibility that the abundance of a modification does not change significantly in a particular cancer setting, but its genomic distribution does. It is probable that a mutation in, or altered expression of, an epigenetic modifier could lead to a reorganization of a modification at key genes without markedly changing its overall abundance. Importantly, the core histones are among the most abundant proteins in eukaryotic cells, with approximately 1×10^8 copies each per cell [83]. As depicted in Figure 1, histone MS data is often visualized in a volcano plot, in which the x-axis value of a point reflects the $\log_2$ ratio of the abundance of a modification across two conditions, and the y-axis represents the $-\log_2$ of the p-value. Statistical significance can be easily calculated, but unfortunately, biological significance cannot. For an abundant modification such as H3K27me2, which is present on approximately 35% of histone H3, a fold-change of 1.1 therefore implies that the status of 3.5×10^6 histones has changed. Such a change could very well be important, but it would not stand out on a volcano plot. Thus, in order to be implicated in a bottom-up MS analysis, a PTM change must be quite profound. We therefore recommend bottom-up histone PTM analysis as a screening method for inferring enzyme function or dysfunction, with the caveat that negative data do not indicate that there is no epigenetic component in the disease process being studied. In cases where there is suspected epigenetic involvement it can be informative to ask how the mutation affects the genomic distribution of the associated mark, even if its abundance is unchanged.

Indeed, when a dysregulated histone modification is identified and somehow correlated with phenotype, investigators often turn to ChIP-seq in order to identify the affected genes and thereby provide a mechanistic link (as outlined in Figure 1). However, interpretation of ChIP-seq data in this context requires that the relationship between the modification and transcription is well understood. This is true for selected single modifications, but the functional output of histones bearing multiple modifications can be more difficult to predict. An alternative or parallel approach that may prove informative in cases where the function of a modification is less clear is the chromatin reader pulldown.

Chromatin readers are proteins with functional domains capable of binding to specific histone modifications – or combinations of modifications – and coupling them to some downstream activity [84]. Much like the MAP3K2 peptide pulldown approach employed by Mazur et al., the reader pulldown approach entails the synthesis of biotinylated histone peptides with or without the modification(s) of interest followed by a pulldown and a comparative analysis of the identified proteins. This method was first utilized in a highly cited paper that reported the binding of TFIID to H3K4me3 [85]. The experimental procedure has been gradually improved since first publication and numerous technical references are available describing both SILAC and label-free implementations [86–88]. The method has also been adapted to identify proteins that bind to specific DNA sequences rather than histone PTMs [89], and this variant was recently significantly improved to utilize TMT labeling methods [90]. With this latest innovation, each bait is loaded onto the beads at 10 different concentrations, so the relative abundance of an interacting partner at each concentration can be converted into an S-shaped binding curve, permitting derivation of

pseudo- K_d values. If this approach can be applied with peptides as baits it will represent a major step forwards in detecting chromatin readers. It is important to note that the peptide pulldown approach has not been specifically applied in a cancer study yet, but it is an important tool for proteomics investigators to be aware of. A schematized workflow is depicted in Figure 2.

Finally, it is possible that mutations in histone modifiers drive oncogenesis in ways that do not depend on histone modification, as is the case with MLL translocations and the recently reported ETS family transcription factor translocations. In these cases, it is most informative to determine how the mutation alters protein-protein interactions rather than histone modification patterns.

3. Complex-centric analysis

3.1. MLL fusions

MLL is a histone H3K4 methyltransferase [91,92]. Over 70% of pediatric leukemias and approximately 10% of adult leukemias bear a translocation involving the *MLL* gene [93]. These translocations typically create fusion genes in which the 5' portion of *MLL* is fused to the 3' portion of another gene, deleting its methyltransferase (SET) domain but retaining several DNA-binding domains. Over 135 translocation partner genes (TPGs) have been described to date, but a few TPGs account for approximately 90% of *MLL* translocations. These are: *ENL*, *ELL*, *AF4*, *AF6*, *AF9*, and *AF10* [94,95]. Thus, MLL fusion partners are not random, they instead appear to be strongly selected for based upon their function. Affinity proteomic characterizations of wild-type MLL, several of its fusion partners, and of chimeric MLL proteins have enabled detailed characterization of how this conversion occurs. For additional reading on the fascinating genetics of MLL fusions we recommend the following reviews: [93,96,97].

In 1998, Slany et al. reported that the transforming capacity of the canonical MLL-ENL fusions required both the DNA binding activity of MLL and the transcriptional effector properties of ENL [98]. This finding, plus the observation that most MLL fusion partners are involved in transcriptional elongation, led to the hypothesis that “a unifying mechanism for the activation of MLL by nuclear fusion partners is the gain of transcriptional effector properties [...]” and “conversion of MLL to a constitutive transcriptional effector may be sufficient to provide a gain of function that promotes leukemogenesis *in vivo*” [99]. In a follow-up study, Mueller and colleagues expressed FLAG-tagged ENL in HEK293 (human embryonic kidney) cells, and then performed IP-MS to identify interaction partners [100]. The list of candidate partners – collectively named EAPs (ENL-associated proteins) – comprised 15 proteins, many of which were previously shown to interact with ENL or its homolog AF9 in 2-hybrid screens. These included the H3K79 methyltransferase DOT1L, the RNA polymerase II C-terminal domain kinase P-TEFb, and the AF4-family transcription factors AF4, LAF4, and AF5q31. Notably, many of the identified interacting proteins can also substitute for ENL as an oncogenic MLL fusion partner, suggesting that the fusion partners are modular, and interchangeably fuse MLL to the complex.

In 2010, an independent study by Yokoyama and colleagues aimed to further characterize the ENL/AF4 containing complex with LC-MS/MS. Similar to the prior study, they purified AF4 complexes from K562 (myelogenous leukemia) cells and identified many of the same proteins including ENL, AF5q31, and P-TEFb. They named this complex “AEP” for AF4 family/ENL family/P-TEFb. They then performed co-IP experiments demonstrating that the fusion proteins MLL-ENL, MLL-AF4, and MLL-AF5q31 aberrantly tether MLL to AEP-like “hybrid complexes.” This was a major step forwards from Mueller et al., because it directly demonstrated that MLL fusions retain many of their canonical protein-protein interactions. Using a series of ChIP-qPCR experiments in two MLL fusion-positive leukemia cell lines, Yokoyama et al. went on to show that these hybrid complexes were aberrantly recruited to and promoted transcriptional elongation at MLL target genes implicated in leukemia such as *CDKN1b*, *CDKN2C*, *HOXA9*, and *MEIS1* [101].

It is tempting to think of the contribution of MLL to MLL-fusion mediated leukemogenesis as merely a DNA-binder, aberrantly tethering its partner and associated transcriptional elongation factors to specific genetic loci. Though this does occur, MLL also mediates important protein-protein interactions that are required for leukemogenesis. The first of these identified was the interaction with Menin, discovered in 2004 by Yokoyama and colleagues in an IP-MS experiment targeted against wild-type MLL. They found that among the identified interacting partners, Menin was the only protein whose knockdown altered MLL-mediated *Hox* gene expression [102]. A later proteomics study revealed that Menin recruits the transcriptional activator LEDGF to MLL target loci, and that this is necessary for both normal and oncogenic MLL-driven transcription [103].

In 2010, Muntean et al. sought to characterize alternative or additional MLL-mediated protein-protein interactions relevant to leukemogenesis [104]. Motivated by a previous report that the CxxC domain and surrounding regions were required for immortalization by MLL fusion proteins [105] they transiently expressed FLAG-tagged portions of this region in HEK293 cells and performed IP-MS, ultimately identifying several subunits of the polymerase associated factor complex (PAFc). These interactions were maintained with both wild-type MLL and the MLL-AF9 fusion protein, and PAFc colocalized with MLL-AF9 at MLL target genes. Importantly, mutations in the PAFc-binding region of MLL prevented cell transformation in 5-fluorouracil primed bone marrow cells without affecting the interaction between the fusion protein and EAP complex members [104].

Several additional proteomics studies have attempted to characterize the biology of MLL fusions in greater detail [106–108]. Earlier this year, a systematic effort was made to comprehensively survey the protein-protein interactions of seven different MLL-fusion proteins [109]. The investigators inserted single, affinity-tagged, copies of each of the seven MLL fusions of interest (-AF1p, -AF4, -AF9, -CBP, -EEN, -ENL, and -GAS7) into isogenic Jurkat cells under the control of a Dox-inducible promoter. They then performed seven independent IP-MS experiments, detecting over 4600 proteins in total. These results were filtered to identify a core group of 128 proteins present in five or more IPs. Members of this core network included usual suspects such as Menin and LEDGF, but also included a number of proteins not previously associated with MLL, chief among which was the histone H3K36 trimethyltransferase SETD2 [109]. Follow-up experiments suggested that SETD2 is

important for both cell transformation and sustained growth in the presence of MLL fusions, highlighting SETD2 as a potential therapeutic target in MLL-associated leukemias. This study represents a valuable resource for MLL investigators and will likely inspire a significant amount of follow-up work by laboratories around the world.

The MLL fusion studies described above are illustrative of the general approach of complex-centric analysis as applied to a fusion protein. They all proceeded from the hypothesis that the biology of the fusion protein can be inferred from its interaction networks, and those of its component parts, as illustrated in Figure 3. By systematically interrogating these networks with proteomics, researchers have identified critical disease mediators and numerous therapeutic targets for these deadly leukemias.

3.2. Additional oncogenic fusions with putative chromatin-modifying activity

Chromosomal translocations fusing the ETS-family transcription factor *ERG* to the serine protease *TMPRSS2* occur in approximately half of prostate cancers (reviewed in [110]). *ERG* is an ETS-family transcription factor that is expressed in endothelial cells and hematopoietic tissue, but not in normal prostate or other epithelium [111]. The *TMPRSS2-ERG* fusion leads to a profound elevation in *ERG* expression and correlates with poor survival, but the mechanism by which it drives oncogenesis has remained somewhat elusive [112,113]. A recent study by Sandoval and colleagues utilized IP-MS to identify proteins that interact with *ERG* in a translocation-positive prostate cancer cell line [114]. This study revealed that *ERG* binds to the SWI/SNF (BAF) ATP-dependent chromatin remodeling complex and can recruit it to ETS target loci across the genome. Other ETS-family transcription factors implicated in prostate cancer including *ETV1*, *ETV4*, and *ETV5* can also bind to SWI/SNF, suggesting that this may be a common mechanism.

This is but one of many examples of the alterations in chromatin remodelers found in cancer (reviewed in [115–117]). A selection of these that feature particularly interesting fusions is depicted in Table 2. However, it is important to note that the alterations in cancer include both mutations and chromosomal translocations. Since the affected proteins are typically components of large, multi-subunit complexes, they likely alter the composition, function, and/or targeting of these complexes. Indeed, this has already been demonstrated for some of the cancer-associated fusion proteins [115,118].

3.3. Pitfalls of affinity proteomics for inferring oncogenic mechanisms of fusions

When advertising the appeal of proteomics, proponents often cite several drawbacks of antibody-based protein or PTM identification techniques. These include antibody cross-reactivity, the opportunity for epitope occlusion, the narrow linear range of quantitation, the need for antibody optimization, and more. Affinity proteomics generally reintroduces these problems into proteomics workflows. Some of these problems can be avoided if target proteins are engineered to be epitope-tagged, but this workaround raises concerns about the effects of the epitope on native protein localization, stability, or protein-protein interactions. This also precludes the analysis of primary tissue, which is especially problematic for researchers interested in the aforementioned translocations, as they mainly occur in rare sarcomas. An additional complication associated with the use of transgenic models is that

they may confound effects of the translocation on native regulation of gene expression, which can be critical to pathogenesis, as with *TMPRSS2-ERG*. Affinity proteomics approaches are therefore most useful when the mutant gene product is known to be translated into a fusion protein, and especially when both partners possess functional domains with known or predicted protein-binding activity.

Finally, like histone PTM analysis, IP-MS typically yields a list of candidates that require functional follow-up. Results from affinity purification experiments can help guide the design of ChIP-seq and/or ChIP-qPCR experiments for specific proteins, but, in isolation, merely reveal what is changing. The next steps are to identify the genomic loci at which the change is occurring, and how this affects downstream transcription. Proteomics is not yet capable of this. However, of note, a recently developed approach termed genomic locus proteomics (“GloPro”) allows for the interrogation of proteins associated with a specific DNA sequence [126]. Therefore, it may soon be possible for proteomics investigators to detect hybrid complex formation at specific disease-associated loci without the use of sequencing technologies.

4. Special Case: Oncohistones

An “oncohistone” is a cancer-associated mutant histone (reviewed in [127]). Oncohistones were first described 2012, when two studies simultaneously reported H3.1K27M, H3.3K27M, and H3.3 G34R/V mutations in pediatric high-grade gliomas [128,129]. Additional oncohistone mutations have since been discovered, including H3.3K27M in posterior fossa ependymoma [130], H3.3G34W in giant cell tumor of bone and osteosarcoma, H3.1K36M in head and neck squamous cell carcinoma (HNSCC), and H3.3 K36M in HNSCC and chondroblastoma [131–134]. These are summarized in Figure 4. Proteomic analyses have supported many of the key studies that are beginning to shed light on the oncogenic mechanisms of oncohistones.

The first mechanistic study of an oncohistone was published in 2013 [135]. Lewis et al. reported that cultured cells and tissue slices from diffuse intrinsic pontine gliomas (DIPGs) expressing H3.3K27M displayed markedly low levels of histone H3K27me_{2/3}. Accordingly, H3.3K27M expression was sufficient to reduce H3K27me₃ levels *in vitro* and *in vivo* via potent inhibition of the EZH2 subunit of PRC2. The study also revealed that H3.3K9M and K36M can likewise inhibit their respective SET domain-containing methyltransferases. The major contribution of MS to this study was the measurement of mutant histone levels relative to wild-type in affected cells. A common misconception is that K27 methylation levels are reduced in K27M-expressing cells simply because lysines are replaced by methionines. In fact, the mutant histone typically represents a minority of total histone H3 (ranging from 3 to 17% in the three DIPG lines tested), supporting dominant negative-type inhibition of K27 methylation [135]. This model was also supported by a concomitant study that primarily used western blotting and ChIP-sequencing to assess mutation-induced changes in the global histone methylation landscape [136].

As predicted in the 2013 study from Lewis et al., dominant negative inhibition of residue-specific methyltransferases has also been observed in tumors expressing H3K36M mutants.

Histone PTM analysis in chondroblastoma cells and chondrocyte models expressing H3.3K36M revealed a global reduction in H3K36me2/3 [137]. MS also supported a later study demonstrating that H3.3K36M is associated with dramatic losses of H3K36me2/3 in HNSCC cells [134].

Unlike H3K27 and K36, G34 is not post-translationally modified, indicating that G34 mutant histones might have a distinct oncogenic mechanism. In fact, in the initial characterization of K27M, Lewis and colleagues reported that expression of G34R or G34V in 293 cells did not induce global changes in histone PTMs, but that the mutant histones displayed lower levels of K36me2/3. Moreover, in an *in vitro* methyltransferase assay, heterotypic mononucleosomes containing the G34 mutants were refractory to methylation by SETD2 [135]. Recently, it was reported that when expressed in HeLa cells, G34L and G34W mutant histones themselves display altered methylation patterns, but do not induce global changes (as measured by western blot) [138].

In addition to altering histone PTMs, mutant histones may alter protein-protein interactions. To determine whether this is the case, Lim and colleagues recently performed an affinity proteomic characterization of GFP-tagged H3.3 G34R and G34W in 293T cells. They observed a modest enrichment for proteins associated with RNA processing, particularly with G34W. Consistent with this finding, subsequent RNA-seq in G34W-expressing tumors revealed defects in alternative splicing, but the potential link between these observations and tumorigenesis remains to be investigated.

Many questions remain about the oncogenic mechanisms of histone mutations. While K27M and K36M are clearly dominant-negative inhibitors of their respective methyltransferases, their effects are more complex than a simple loss-of-function. Although they induce dramatic losses in their respective methylation marks, the marks do not reach zero. In fact cells expressing H3.3K27M display aberrantly high levels of K27me3 at a subset of genomic loci including key tumor suppressors [136,139]. Similarly, ChIP-seq of H3K36me2 in chondrocytes expressing H3.3K36M revealed over 7000 peaks that were not present in cells with wild-type H3.3 [137]. The fact that loss-of-function mutations in the associated histone methyltransferases have never been observed in the same cancers that possess oncohistone mutations provides additional support for the hypothesis that they operate by distinct mechanisms [140].

Notably, though the associated mechanisms are incompletely understood, a common theme that has emerged is that cells expressing oncohistones are refractory to differentiation [127]. This observation is rather consistent with Jay Roth's 1965 hypothesis that mutations "in the gene or gene system related to histone control" could cause "a reversion of the cell to a less differentiated type" [37]. In summary, mutant histones can induce global (K → M) or focal (G34X) changes to the histone methylation landscape, and they may recruit specific chromatin-interacting partners. The extent to which each of these effects contribute to oncogenesis remains to be determined.

5. Conclusion

The observation that histones from cancer cells tend to display abnormal biophysical properties preceded the discovery of histone post-translational modifications by ten years. Mass spectrometry was not required for either discovery, but it has been invaluable in unifying them and in facilitating the characterization of the accompanying mechanisms. When applied in a systematic fashion, proteomic analyses of histone PTMs have identified recurrent alterations that occur in cancer and can predict sensitivity to targeted therapeutics. When applied in specific cases, such analyses have enabled the functional characterization of a recurrent driver mutation and the identification of a novel mechanism of tumor suppressor repression. Nevertheless it is important to acknowledge that the histone modification-centric approach is limited in that it does not identify potential nonhistone substrates that might contribute to cancer phenotype. As currently implemented, it simply reports the global level of individual marks without providing information about their genomic localization. Histone PTM analysis can often suggest directions for downstream experimentation with alternative technologies such as ChIP-seq, but proteomic identification of specific chromatin readers is also a potentially informative follow-up approach.

The other main strength of proteomic analysis with respect to cancer epigenetics research is the ability to identify unusual protein-protein interactions that may occur as a result of alterations in epigenetic modifiers. This approach has been successfully used to identify the interacting partners of MLL, many of its individual translocation partners, and of a number of chimeric MLL fusion proteins. Collectively these studies have clarified the mechanism by which a subset of MLL fusion proteins lead to aberrant transcription, and supported the development of inhibitors to DOT1L, Menin, and – in the wake of Skucha et al. – perhaps SETD2 for leukemia therapy. The recent report of ETS-family transcription factors/BAF complex formation in prostate cancer suggests that proteomic analysis will be valuable in the characterization of additional oncogenic chromosomal translocations affecting chromatin modifiers that as of yet remain poorly understood. Moreover, this approach need not be limited to fusion proteins. Similar approaches could address mutations or other more subtle alterations affecting the function of epigenetic regulators, thereby offering insight into a broader array of human tumors.

6. Expert commentary

This review has been limited in scope to the analysis of histone PTMs and protein complexes as they relate to cancer epigenetics. A common format among the modification-centric studies presented here has been the identification of a single histone PTM attributable to a single dysregulated enzyme in a single cancer type. These studies are generally followed by experiments aimed at determining how this alteration affects transcription, which is, of course, done without the use of MS. This has been an informative workflow, but mass spectrometers are immensely versatile instruments, and to limit their contribution to this would be an underutilization. One major unexplored area is the potential existence of dysregulated histone PTM crosstalk in cancer. Middle-down and/or top-down analyses, which allow for the quantification of combinatorial histone marks, have not yet been systematically applied in any cancer study. We suspect that some cancer-associated

mutations might manifest in altered histone PTM combinations, and that these might mediate the repression or expression of key genes. MS will be essential first to test this hypothesis, but if unusual histone PTM combinations are observed, it will also be important to help decipher the effects of these combinations on reader recruitment.

Until very recently, proteomics strategies for identifying chromatin readers were largely qualitative. The recent report by Makowski et al. adds a quantitative dimension to these assays, allowing for a more accurate description of reader behavior [90]. If applied to combinatorially modified peptides, this approach should allow for the identification of “master readers” that bind to specific marks regardless of PTM coexistence, and of readers whose binding is more sensitive and therefore perhaps more tightly regulated. It will be particularly exciting if it is possible to unify this general experimental design with more physiologically relevant substrates, such as the recently described “designer nucleosomes” which can be engineered to bear virtually any combination of PTMs [141].

Both combinatorial histone PTM analysis and improved chromatin reader profiling could also be applied to oncohistones. Bottom-up and western blot analysis of global PTM levels in oncohistone-expressing cells has suggested a simple dichotomy in which $K \rightarrow M$ mutations work in trans via dominant negative mechanisms and $G \rightarrow X$ mutations work by disrupting SETD2-dependent methylation in cis, but this is likely an oversimplification. In bottom-up workflows, all oncohistones studied to date (K27M, G34X, and K36M) fall on the same tryptic peptide, which means that it is still unknown if they bear unusual modification patterns at distal sites such as K4 and K9. It is also unknown whether they recruit inappropriate interacting partners and how this might contribute to their cellular effects.

7. Five-year view

We anticipate that many groups will turn to FFPE samples to access the histone PTM profile of primary tumors. This will be especially informative in the study of rare, poorly understood tumors known to have genetic alterations in epigenetic regulators. In addition, it should soon become feasible to analyze histones from cancer cells in a middle-down fashion in a highly robust and reproducible way. There may first be analyses in the spirit of Cruft and Fraga, aiming to identify recurrent PTM combinations that occur across a broad range of cancer types. These would be followed by more focused reports of combinatorial PTM signatures that occur in specific cancers, perhaps associated with expression of a particular mutant histone or modifier. At the same time, there may be increased adoption of chromatin reader profiling techniques, potentially investigating combinatorial marks either on peptides or on designer nucleosomes. There may also be a movement away from the MLL fusions, many of which are now well understood, and an increased focus on members of the long list of additional chromosomal translocations that remain poorly characterized. We also anticipate that, as genomic locus proteomic approaches mature, they too may become powerful tools to study epigenetic dysregulation in cancer.

Acknowledgments

Funding

D. Marchione has received funding from the US National Institutes of health, grants: NIH TL1TR001880 and NIH T32GM008275. J. Wojcik has received funding from the University of Pennsylvania, Abramson Cancer Center, Paul Calabresi K12 Career Development Award for Clinical Oncology (5K12CA076931). B.A. Garcia has received funding from the US National Institutes of health grants: NIH GM110174, NIH CA196539; and a Robert Arceci Scholar Award from the Leukemia and Lymphoma Society.

References

Papers of special note have been highlighted as:

* of interest

** of considerable interest

- [1]. Feinberg AP, Koldobskiy MA, Göndör A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat. Rev. Genet* 2016;17:284–299. [PubMed: 26972587]
- [2]. Shen L, Shi Q, Wang W. Double agents: genes with both oncogenic and tumor-suppressor functions. *Oncogenesis* 2018;7:25. [PubMed: 29540752]
- [3]. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat. Med* 2004;10:789–799. [PubMed: 15286780] *Provides a theoretical framework for understanding the genetic basis of cancer. Also acknowledges the importance of gene expression changes that occur in the absence of mutations
- [4]. Landau DA, Clement K, Ziller MJ, et al. Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. *Cancer Cell* 2014;26:813–825. [PubMed: 25490447]
- [5]. You JS, Jones PA. Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell* 2012;22:9–20. [PubMed: 22789535]
- [6]. Huang H, Sabari BR, Garcia BA, et al. SnapShot: histone modifications. *Cell* 2014;159:458–458.e1. [PubMed: 25303536]
- [7]. Strahl BD, Allis CD. The language of covalent histone modifications. *Nature* 2000;403:41–45. [PubMed: 10638745] *Presents the idea that histone modifications function like a language regulating gene expression
- [8]. Jenuwein T, Allis CD. Translating the histone code. *Science* 2001;293:1074–1080. [PubMed: 11498575]
- [9]. Berger SL. The complex language of chromatin regulation during transcription. *Nature* 2007;447:407–412. [PubMed: 17522673]
- [10]. Barski A, Cuddapah S, Cui K, et al. High-resolution profiling of histone methylations in the human genome. *Cell* 2007;129:823–837. [PubMed: 17512414] *Generated ChIP-seq profiles for 20 histone PTMs and assessed correlations between their localization and gene expression
- [11]. Lehnertz B, Ueda Y, Derijck AAHA, et al. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol* 2003;13:1192–1200. [PubMed: 12867029]
- [12]. Francis NJ, Kingston RE, Woodcock CL. Chromatin compaction by a polycomb group protein complex. *Science* 2004;306:1574–1577. [PubMed: 15567868]
- [13]. Creighton MP, Cheng AW, Welstead GG, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U. S. A* 2010;107:21931–21936. [PubMed: 21106759]
- [14]. Tan M, Luo H, Lee S, et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 2011;146:1016–1028. [PubMed: 21925322]
- [15]. Dai L, Peng C, Montellier E, et al. Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. *Nat. Chem. Biol* 2014;10:365–370. [PubMed: 24681537]
- [16]. Huang H, Zhang D, Wang Y, et al. Lysine benzoylation is a histone mark regulated by SIRT2. *Nat. Commun* 2018;9:3374. [PubMed: 30154464]
- [17]. Sidoli S, Schwämmle V, Ruminowicz C, et al. Middle-down hybrid chromatography/tandem mass spectrometry workflow for characterization of combinatorial post-translational

modifications in histones. *Proteomics* 2014;14:2200–2211. [PubMed: 25073878] *Presented a robust middle-down strategy for profiling combinatorial histone PTMs

- [18]. Schwämmle V, Sidoli S, Ruminowicz C, et al. Systems Level Analysis of Histone H3 Post-translational Modifications (PTMs) Reveals Features of PTM Crosstalk in Chromatin Regulation. *Mol. Cell. Proteomics* 2016;15:2715–2729. [PubMed: 27302890]
- [19]. Zheng Y, Fornelli L, Compton PD, et al. Unabridged Analysis of Human Histone H3 by Differential Top-Down Mass Spectrometry Reveals Hypermethylated Proteoforms from MMSET/NSD2 Overexpression. *Mol. Cell. Proteomics* 2016;15:776–790. [PubMed: 26272979]
- [20]. Wang T, Holt MV, Young NL. Early butyrate induced acetylation of histone H4 is proteoform specific and linked to methylation state. *Epigenetics* 2018;1–17. *Used top-down to reveal histone PTM crosstalk that affects the response to drug treatment
- [21]. Cristobal A, Marino F, Post H, et al. Toward an Optimized Workflow for Middle-Down Proteomics. *Anal. Chem* 2017;89:3318–3325. [PubMed: 28233997]
- [22]. Sidoli S, Garcia BA. Middle-down proteomics: a still unexploited resource for chromatin biology. *Expert Rev. Proteomics* 2017;14:617–626. [PubMed: 28649883]
- [23]. Sidoli S, Vandamme J, Salcini AE, et al. Dynamic changes of histone H3 marks during *Caenorhabditis elegans* lifecycle revealed by middle-down proteomics. *Proteomics* 2016;16:459–464. [PubMed: 26508544]
- [24]. Schröder CU, Ziemianowicz DS, Merx K, et al. Simultaneous Proteoform Analysis of Histones H3 and H4 with a Simplified Middle-Down Proteomics Method. *Anal. Chem* 2018;90:3083–3090. [PubMed: 29405698]
- [25]. Catherman AD, Skinner OS, Kelleher NL. Top Down proteomics: facts and perspectives. *Biochem. Biophys. Res. Commun* 2014;445:683–693. [PubMed: 24556311]
- [26]. Dunham WH, Mullin M, Gingras A-C. Affinity-purification coupled to mass spectrometry: basic principles and strategies. *Proteomics* 2012;12:1576–1590. [PubMed: 22611051]
- [27]. Yates JR 3rd. The revolution and evolution of shotgun proteomics for large-scale proteome analysis. *J. Am. Chem. Soc* 2013;135:1629–1640. [PubMed: 23294060]
- [28]. Aebersold R, Mann M. Mass-spectrometric exploration of proteome structure and function. *Nature* 2016;537:347–355. [PubMed: 27629641]
- [29]. Huang H, Lin S, Garcia BA, et al. Quantitative proteomic analysis of histone modifications. *Chem. Rev* 2015;115:2376–2418. [PubMed: 25688442]
- [30]. Zheng Y, Huang X, Kelleher NL. Epiroteomics: quantitative analysis of histone marks and codes by mass spectrometry. *Curr. Opin. Chem. Biol* 2016;33:142–150. [PubMed: 27371874]
- [31]. Janssen KA, Sidoli S, Garcia BA. Recent Achievements in Characterizing the Histone Code and Approaches to Integrating Epigenomics and Systems Biology. *Methods Enzymol* 2017;586:359–378. [PubMed: 28137571]
- [32]. Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. *Cell* 2012;150:12–27. [PubMed: 22770212]
- [33]. Flavahan WA, Gaskell E, Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. *Science [Internet]* 2017;357 Available from: 10.1126/science.aal2380.
- [34]. Stedman E, Stedman E. Cell specificity of histones. *Nature* 1950;166:780–781. [PubMed: 14780242] **First statement of the hypothesis that histones are related to cell identity
- [35]. Cruft HJ, Mauritzen CM, Stedman E. Abnormal properties of histones from malignant cells. *Nature* 1954;174:580–585. [PubMed: 13203581] **Discovery that histones from cancer cells are somehow different from those of normal cells
- [36]. Allfrey VG, Faulkner R, Mirsky AE. ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. *Proc. Natl. Acad. Sci. U. S. A* 1964;51:786–794. [PubMed: 14172992] **Discovery that histones are modified and contribute to regulation of gene expression
- [37]. Roth JS. Histones in development, growth and cancer. *Nature* 1965;207:599–601. [PubMed: 5883632] **Prediction of epigenetic cancer mechanisms
- [38]. Fraga MF, Ballestar E, Villar-Garea A, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat. Genet* 2005;37:391–400.

[PubMed: 15765097] *First use of mass spectrometry to analyze the relationship between histone PTMs and cancer

- [39]. Cuomo A, Moretti S, Minucci S, et al. SILAC-based proteomic analysis to dissect the “histone modification signature” of human breast cancer cells. *Amino Acids* 2011;41:387–399. [PubMed: 20617350]
- [40]. Leroy G, Dimaggio PA, Chan EY, et al. A quantitative atlas of histone modification signatures from human cancer cells. *Epigenetics Chromatin* 2013;6:20. [PubMed: 23826629]
- [41]. Kaniskan HÜ, Konze KD, Jin J. Selective inhibitors of protein methyltransferases. *J. Med. Chem* 2015;58:1596–1629. [PubMed: 25406853]
- [42]. Barsyte-Lovejoy D, Li F, Oudhoff MJ, et al. (R)-PFI-2 is a potent and selective inhibitor of SETD7 methyltransferase activity in cells. *Proc. Natl. Acad. Sci. U. S. A* 2014;111:12853–12858. [PubMed: 25136132]
- [43]. Cao F, Townsend EC, Karatas H, et al. Targeting MLL1 H3K4 methyltransferase activity in mixed-lineage leukemia. *Mol. Cell* 2014;53:247–261. [PubMed: 24389101]
- [44]. Xiong Y, Li F, Babault N, et al. Discovery of Potent and Selective Inhibitors for G9a-Like Protein (GLP) Lysine Methyltransferase. *J. Med. Chem* 2017;60:1876–1891. [PubMed: 28135087]
- [45]. Kuntz KW, Campbell JE, Keilhack H, et al. The Importance of Being Me: Magic Methyls, Methyltransferase Inhibitors, and the Discovery of Tazemetostat. *J. Med. Chem* 2016;59:1556–1564. [PubMed: 26769278]
- [46]. Italiano A, Soria J-C, Toulmonde M, et al. Tazemetostat, an EZH2 inhibitor, in relapsed or refractory B-cell non-Hodgkin lymphoma and advanced solid tumours: a first-in-human, open-label, phase 1 study. *Lancet Oncol* 2018;19:649–659. [PubMed: 29650362]
- [47]. Verma SK, Tian X, LaFrance LV, et al. Identification of Potent, Selective, Cell-Active Inhibitors of the Histone Lysine Methyltransferase EZH2. *ACS Med. Chem. Lett* 2012;3:1091–1096. [PubMed: 24900432]
- [48]. Konze KD, Ma A, Li F, et al. An orally bioavailable chemical probe of the Lysine Methyltransferases EZH2 and EZH1. *ACS Chem. Biol* 2013;8:1324–1334. [PubMed: 23614352]
- [49]. Zheng W, Ibáñez G, Wu H, et al. Sinefungin derivatives as inhibitors and structure probes of protein lysine methyltransferase SETD2. *J. Am. Chem. Soc* 2012;134:18004–18014. [PubMed: 23043551]
- [50]. Daigle SR, Olhava EJ, Therkelsen CA, et al. Potent inhibition of DOT1L as treatment of MLL-fusion leukemia. *Blood* 2013;122:1017–1025. [PubMed: 23801631]
- [51]. Ma A, Yu W, Li F, et al. Discovery of a selective, substrate-competitive inhibitor of the lysine methyltransferase SETD8. *J. Med. Chem* 2014;57:6822–6833. [PubMed: 25032507]
- [52]. Bromberg KD, Mitchell TRH, Upadhyay AK, et al. The SUV4–20 inhibitor A-196 verifies a role for epigenetics in genomic integrity. *Nat. Chem. Biol* 2017;13:317–324. [PubMed: 28114273]
- [53]. Noberini R, Osti D, Miccolo C, et al. Extensive and systematic rewiring of histone post-translational modifications in cancer model systems. *Nucleic Acids Res* 2018;46:3817–3832. [PubMed: 29618087]
- [54]. Noberini R, Uggetti A, Pruneri G, et al. Pathology Tissue-quantitative Mass Spectrometry Analysis to Profile Histone Post-translational Modification Patterns in Patient Samples. *Mol. Cell. Proteomics* 2016;15:866–877. [PubMed: 26463340]
- [55]. Noberini R, Longuespée R, Richichi C, et al. PAT-H-MS coupled with laser microdissection to study histone post-translational modifications in selected cell populations from pathology samples. *Clin. Epigenetics* 2017;9:69. [PubMed: 28702092]
- [56]. Chesi M, Nardini E, Lim RS, et al. The t(4;14) translocation in myeloma dysregulates both FGFR3 and a novel gene, MMSET, resulting in IgH/MMSET hybrid transcripts. *Blood* 1998;92:3025–3034. [PubMed: 9787135]
- [57]. Keats JJ, Reiman T, Maxwell CA, et al. In multiple myeloma, t(4;14)(p16;q32) is an adverse prognostic factor irrespective of FGFR3 expression. *Blood* 2003;101:1520–1529. [PubMed: 12393535]
- [58]. Kim J-Y, Kee HJ, Choe N-W, et al. Multiple-myeloma-related WHSC1/MMSET isoform RE-IIBP is a histone methyltransferase with transcriptional repression activity. *Mol. Cell. Biol* 2008;28:2023–2034. [PubMed: 18172012]

- [59]. Marango J, Shimoyama M, Nishio H, et al. The MMSET protein is a histone methyltransferase with characteristics of a transcriptional corepressor. *Blood* 2008;111:3145–3154. [PubMed: 18156491]
- [60]. Li Y, Trojer P, Xu C-F, et al. The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate. *J. Biol. Chem* 2009;284:34283–34295. [PubMed: 19808676]
- [61]. Nimura K, Ura K, Shiratori H, et al. A histone H3 lysine 36 trimethyltransferase links Nkx2–5 to Wolf-Hirschhorn syndrome. *Nature* 2009;460:287–291. [PubMed: 19483677]
- [62]. Martinez-Garcia E, Popovic R, Min D-J, et al. The MMSET histone methyl transferase switches global histone methylation and alters gene expression in t(4;14) multiple myeloma cells. *Blood* 2011;117:211–220. [PubMed: 20974671] *Describes the oncogenic mechanism of NSD2 in multiple myeloma. Presents a general workflow for linking aberrant histone modifier activity to cancer phenotype
- [63]. Pei H, Zhang L, Luo K, et al. MMSET regulates histone H4K20 methylation and 53BP1 accumulation at DNA damage sites. *Nature* 2011;470:124–128. [PubMed: 21293379]
- [64]. Kuo AJ, Cheung P, Chen K, et al. NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. *Mol. Cell* 2011;44:609–620. [PubMed: 22099308]
- [65]. Michishita E, Park JY, Burneskis JM, et al. Evolutionarily conserved and nonconserved cellular localizations and functions of human SIRT proteins. *Mol. Biol. Cell* 2005;16:4623–4635. [PubMed: 16079181]
- [66]. Barber MF, Michishita-Kioi E, Xi Y, et al. SIRT7 links H3K18 deacetylation to maintenance of oncogenic transformation. *Nature* 2012;487:114–118. [PubMed: 22722849]
- [67]. Wang Z, Zang C, Rosenfeld JA, et al. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet* 2008;40:897–903. [PubMed: 18552846]
- [68]. Seligson DB, Horvath S, McBrien MA, et al. Global levels of histone modifications predict prognosis in different cancers. *Am. J. Pathol* 2009;174:1619–1628. [PubMed: 19349354]
- [69]. Manuyakorn A, Paulus R, Farrell J, et al. Cellular histone modification patterns predict prognosis and treatment response in resectable pancreatic adenocarcinoma: results from RTOG 9704. *J. Clin. Oncol* 2010;28:1358–1365. [PubMed: 20142597]
- [70]. Hamamoto R, Furukawa Y, Morita M, et al. SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat. Cell Biol* 2004;6:731–740. [PubMed: 15235609]
- [71]. Hamamoto R, Silva FP, Tsuge M, et al. Enhanced SMYD3 expression is essential for the growth of breast cancer cells. *Cancer Sci* 2006;97:113–118. [PubMed: 16441421]
- [72]. Van Aller GS, Reynoird N, Barbash O, et al. Smyd3 regulates cancer cell phenotypes and catalyzes histone H4 lysine 5 methylation. *Epigenetics* 2012;7:340–343. [PubMed: 22419068]
- [73]. Foreman KW, Brown M, Park F, et al. Structural and functional profiling of the human histone methyltransferase SMYD3. *PLoS One* 2011;6:e22290. [PubMed: 21779408]
- [74]. Mazur PK, Reynoird N, Khatri P, et al. SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer. *Nature* 2014;510:283–287. [PubMed: 24847881]
- [75]. Yao Y-L, Yang W-M. Beyond Histone and Deacetylase: An Overview of Cytoplasmic Histone Deacetylases and Their Nonhistone Substrates. *Biomed Res. Int.* [Internet] 2010 [cited 2018 Sep 11];2011. Available from: <https://www.hindawi.com/journals/bmri/2011/146493/>.
- [76]. Martínez-Redondo P, Vaquero A. The diversity of histone versus nonhistone sirtuin substrates. *Genes Cancer* 2013;4:148–163. [PubMed: 24020006] *Discovery and characterization of bivalent chromatin
- [77]. Weigt D, Hopf C, Médard G. Studying epigenetic complexes and their inhibitors with the proteomics toolbox. *Clin. Epigenetics* 2016;8:76. [PubMed: 27437033]
- [78]. Bernstein BE, Mikkelsen TS, Xie X, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 2006;125:315–326. [PubMed: 16630819]
- [79]. Voigt P, LeRoy G, Drury WJ 3rd, et al. Asymmetrically modified nucleosomes. *Cell* 2012;151:181–193. [PubMed: 23021224]
- [80]. Schmitges FW, Prusty AB, Faty M, et al. Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol. Cell* 2011;42:330–341. [PubMed: 21549310]

- [81]. Yuan W, Xu M, Huang C, et al. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem* 2011;286:7983–7989. [PubMed: 21239496]
- [82]. Cai L, Rothbart SB, Lu R, et al. An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. *Mol. Cell* 2013;49:571–582. [PubMed: 23273982]
- [83]. Schwanhäusser B, Busse D, Li N, et al. Global quantification of mammalian gene expression control. *Nature* 2011;473:337–342. [PubMed: 21593866]
- [84]. Taverna SD, Li H, Ruthenburg AJ, et al. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol* 2007;14:1025–1040. [PubMed: 17984965] *Review of the concept of chromatin readers
- [85]. Vermeulen M, Mulder KW, Denissov S, et al. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* 2007;131:58–69. [PubMed: 17884155] *First use of proteomics for screening chromatin readers
- [86]. Vermeulen M, Eberl HC, Matarese F, et al. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 2010;142:967–980. [PubMed: 20850016]
- [87]. Vermeulen M Identifying chromatin readers using a SILAC-based histone peptide pull-down approach. *Methods Enzymol* 2012;512:137–160. [PubMed: 22910206]
- [88]. Eberl HC, Spruijt CG, Kelstrup CD, et al. A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics. *Mol. Cell* 2013;49:368–378. [PubMed: 23201125] *Improved workflow for proteomic screening of DNA readers
- [89]. Spruijt CG, Baymaz HI, Vermeulen M. Identifying Specific Protein–DNA Interactions Using SILAC-Based Quantitative Proteomics. In: Bina M, editor. *Gene Regulation: Methods and Protocols* Totowa, NJ: Humana Press; 2013 p. 137–157.
- [90]. Makowski MM, Gräwe C, Foster BM, et al. Global profiling of protein–DNA and protein–nucleosome binding affinities using quantitative mass spectrometry. *Nat. Commun* 2018;9:1653. [PubMed: 29695722]
- [91]. Milne TA, Briggs SD, Brock HW, et al. MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol. Cell* 2002;10:1107–1117. [PubMed: 12453418]
- [92]. Nakamura T, Mori T, Tada S, et al. ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol. Cell* 2002;10:1119–1128. [PubMed: 12453419]
- [93]. Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. *Nat. Rev. Cancer* 2007;7:823–833. [PubMed: 17957188]
- [94]. Meyer C, Schneider B, Jakob S, et al. The MLL recombinome of acute leukemias. *Leukemia* 2006;20:777–784. [PubMed: 16511515]
- [95]. Meyer C, Burmeister T, Gröger D, et al. The MLL recombinome of acute leukemias in 2017. *Leukemia* 2018;32:273–284. [PubMed: 28701730]
- [96]. Marschalek R Mechanisms of leukemogenesis by MLL fusion proteins. *Br. J. Haematol* 2011;152:141–154. [PubMed: 21118195]
- [97]. Yokoyama A Transcriptional activation by MLL fusion proteins in leukemogenesis. *Exp. Hematol* 2017;46:21–30. [PubMed: 27865805]
- [98]. Slany RK, Lavau C, Cleary ML. The oncogenic capacity of HRX-ENL requires the transcriptional transactivation activity of ENL and the DNA binding motifs of HRX. *Mol. Cell. Biol* 1998;18:122–129. [PubMed: 9418860]
- [99]. Ayton PM, Cleary ML. Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. *Oncogene* 2001;20:5695–5707. [PubMed: 11607819]
- [100]. Mueller D, Bach C, Zeisig D, et al. A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood* 2007;110:4445–4454. [PubMed: 17855633] *First characterization of a protein-protein interaction network associated with an MLL fusion partner
- [101]. Yokoyama A, Lin M, Naresh A, et al. A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription. *Cancer Cell* 2010;17:198–212. [PubMed: 20153263] *First proteomic study confirming hybrid complex assembly

- [102]. Yokoyama A, Wang Z, Wysocka J, et al. Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol. Cell. Biol* 2004;24:5639–5649. [PubMed: 15199122]
- [103]. Yokoyama A, Cleary ML. Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell* 2008;14:36–46. [PubMed: 18598942]
- [104]. Muntean AG, Tan J, Sitwala K, et al. The PAF complex synergizes with MLL fusion proteins at HOX loci to promote leukemogenesis. *Cancer Cell* 2010;17:609–621. [PubMed: 20541477]
- [105]. Bach C, Mueller D, Buhl S, et al. Alterations of the CxxC domain preclude oncogenic activation of mixed-lineage leukemia 2. *Oncogene* 2009;28:815–823. [PubMed: 19060922]
- [106]. Monroe SC, Jo SY, Sanders DS, et al. MLL-AF9 and MLL-ENL alter the dynamic association of transcriptional regulators with genes critical for leukemia. *Exp. Hematol* 2011;39:77–86.e1–e5. [PubMed: 20854876]
- [107]. Okuda H, Kanai A, Ito S, et al. AF4 uses the SL1 components of RNAP1 machinery to initiate MLL fusion- and AEP-dependent transcription. *Nat. Commun* 2015;6:8869. [PubMed: 26593443] *Comprehensive analysis of MLL interaction partners, identifies a core network of proteins associated with MLL fusions
- [108]. Okuda H, Stanojevic B, Kanai A, et al. Cooperative gene activation by AF4 and DOT1L drives MLL-rearranged leukemia. *J. Clin. Invest* 2017;127:1918–1931. [PubMed: 28394257]
- [109]. Skucha A, Ebner J, Schmöllerl J, et al. MLL-fusion-driven leukemia requires SETD2 to safeguard genomic integrity. *Nat. Commun* 2018;9:1983. [PubMed: 29777171]
- [110]. Clark JP, Cooper CS. ETS gene fusions in prostate cancer. *Nat. Rev. Urol* 2009;6:429–439. [PubMed: 19657377]
- [111]. Mohamed AA, Tan S-H, Mikhalkovich N, et al. Ets family protein, erg expression in developing and adult mouse tissues by a highly specific monoclonal antibody. *J. Cancer* 2010;1:197–208. [PubMed: 21060730]
- [112]. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 2005;310:644–648. [PubMed: 16254181]
- [113]. Hägglöf C, Hammarsten P, Strömvall K, et al. TMPRSS2-ERG expression predicts prostate cancer survival and associates with stromal biomarkers. *PLoS One* 2014;9:e86824. [PubMed: 24505269]
- [114]. Sandoval GJ, Pulice JL, Pakula H, et al. Binding of TMPRSS2-ERG to BAF Chromatin Remodeling Complexes Mediates Prostate Oncogenesis. *Mol. Cell* 2018;71:554–566.e7. [PubMed: 30078722]
- [115]. Kadoch C, Crabtree GR. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci Adv* 2015;1:e1500447. [PubMed: 26601204]
- [116]. Wojcik J, Cooper K. Epigenetic Alterations in Bone and Soft Tissue Tumors. *Adv. Anat. Pathol* 2017;24:362–371. [PubMed: 28885261] *Further explores recurrent alterations in epigenetic regulators as they relate to bone and soft tissue tumors
- [117]. Shen C, Vakoc CR. Gain-of-function mutation of chromatin regulators as a tumorigenic mechanism and an opportunity for therapeutic intervention. *Curr. Opin. Oncol* 2015;27:57–63. [PubMed: 25402979]
- [118]. Banito A, Li X, Laporte AN, et al. The SS18-SSX Oncoprotein Hijacks KDM2B-PRC1.1 to Drive Synovial Sarcoma. *Cancer Cell* 2018;33:527–541.e8. [PubMed: 29502955]
- [119]. French CA, Miyoshi I, Kubonishi I, et al. BRD4-NUT fusion oncogene: a novel mechanism in aggressive carcinoma. *Cancer Res* 2003;63:304–307. [PubMed: 12543779]
- [120]. French CA, Ramirez CL, Kolmakova J, et al. BRD-NUT oncoproteins: a family of closely related nuclear proteins that block epithelial differentiation and maintain the growth of carcinoma cells. *Oncogene* 2008;27:2237–2242. [PubMed: 17934517]
- [121]. Gebre-Medhin S, Nord KH, Möller E, et al. Recurrent rearrangement of the PHF1 gene in ossifying fibromyxoid tumors. *Am. J. Pathol* 2012;181:1069–1077. [PubMed: 22796436]
- [122]. Koontz JJ, Soreng AL, Nucci M, et al. Frequent fusion of the JAZF1 and JJAZ1 genes in endometrial stromal tumors. *Proc. Natl. Acad. Sci. U. S. A* 2001;98:6348–6353. [PubMed: 11371647]

- [123]. Kao Y-C, Sung Y-S, Zhang L, et al. Recurrent BCOR Internal Tandem Duplication and YWHAE-NUTM2B Fusions in Soft Tissue Undifferentiated Round Cell Sarcoma of Infancy: Overlapping Genetic Features With Clear Cell Sarcoma of Kidney. *Am. J. Surg. Pathol* 2016;40:1009–1020. [PubMed: 26945340]
- [124]. Karlsson J, Valind A, Gisselsson D. BCOR internal tandem duplication and YWHAE-NUTM2B/E fusion are mutually exclusive events in clear cell sarcoma of the kidney. *Genes Chromosomes Cancer* 2016;55:120–123. [PubMed: 26493387]
- [125]. de Rooij JDE, van den Heuvel-Eibrink MM, Kollen WJW, et al. Recurrent translocation t(10;17) (p15;q21) in minimally differentiated acute myeloid leukemia results in ZMYND11/MBTD1 fusion. *Genes Chromosomes Cancer* 2016;55:237–241. [PubMed: 26608508]
- [126]. Myers SA, Wright J, Peckner R, et al. Discovery of proteins associated with a predefined genomic locus via dCas9-APEX-mediated proximity labeling. *Nat. Methods* [Internet] 2018; Available from: 10.1038/s41592-018-0007-1.*Presents a technique that may be applicable to cancer epigenetics research in the future
- [127]. Weinberg DN, Allis CD, Lu C. Oncogenic Mechanisms of Histone H3 Mutations. *Cold Spring Harb. Perspect. Med.* [Internet] 2017;7 Available from: 10.1101/cshperspect.a026443.*Review of the current understanding of oncohistones
- [128]. Schwartzentruber J, Korshunov A, Liu X-Y, et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* 2012;482:226–231. [PubMed: 22286061]
- [129]. Wu G, Broniscer A, McEachron TA, et al. Somatic Histone H3 Alterations in Paediatric Diffuse Intrinsic Pontine Gliomas and Non-Brainstem Glioblastomas. *Nat. Genet* 2012;
- [130]. Pajtler KW, Wen J, Sill M, et al. Molecular heterogeneity and CXorf67 alterations in posterior fossa group A (PFA) ependymomas. *Acta Neuropathol* 2018;136:211–226. [PubMed: 29909548]
- [131]. Behjati S, Tarpey PS, Presneau N, et al. Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. *Nat. Genet* 2013;45:1479–1482. [PubMed: 24162739]
- [132]. Joseph CG, Hwang H, Jiao Y, et al. Exomic analysis of myxoid liposarcomas, synovial sarcomas, and osteosarcomas. *Genes Chromosomes Cancer* 2014;53:15–24. [PubMed: 24190505]
- [133]. Sarungbam J, Agaram N, Hwang S, et al. Symplastic/pseudoanaplastic giant cell tumor of the bone. *Skeletal Radiol* 2016;45:929–935. [PubMed: 27020452]
- [134]. Papillon-Cavanagh S, Lu C, Gayden T, et al. Impaired H3K36 methylation defines a subset of head and neck squamous cell carcinomas. *Nat. Genet* 2017;49:180–185. [PubMed: 28067913]
- [135]. Lewis PW, Müller MM, Koletsky MS, et al. Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science* 2013;340:857–861. [PubMed: 23539183]
*First mechanistic study of an oncohistone
- [136]. Chan K-M, Fang D, Gan H, et al. The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. *Genes Dev* 2013;27:985–990. [PubMed: 23603901]
- [137]. Fang D, Gan H, Lee J-H, et al. The histone H3.3K36M mutation reprograms the epigenome of chondroblastomas. *Science* 2016;352:1344–1348. [PubMed: 27229140]
- [138]. Shi L, Shi J, Shi X, et al. Histone H3.3 G34 Mutations Alter Histone H3K36 and H3K27 Methylation In Cis. *J. Mol. Biol* 2018;430:1562–1565. [PubMed: 29689253]
- [139]. Mohammad F, Weissmann S, Leblanc B, et al. EZH2 is a potential therapeutic target for H3K27M-mutant pediatric gliomas. *Nat. Med* 2017;23:483–492. [PubMed: 28263309]
- [140]. Lewis PW, Allis CD. Poisoning the “histone code” in pediatric gliomagenesis. *Cell Cycle* 2013;12:3241–3242. [PubMed: 24036540]
- [141]. Dann GP, Liszczak GP, Bagert JD, et al. ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference. *Nature* 2017;548:607–611. [PubMed: 28767641] *Presents a library of designer nucleosomes that may facilitate cancer epigenetics research in the future

8.**Key issues**

- Cancer cell growth is often driven by epigenetic mechanisms
- Proteomic analysis of histone modifications in cancer cells can suggest therapeutic vulnerabilities and provide mechanistic insights into oncogenesis, especially when coupled with techniques like ChIP-sequencing
- Chromatin reader profiling provides another avenue for determining the functional outcome of altered histone modifications
- Affinity proteomics can identify alterations to protein-protein interactions that occur as a result of mutations
- Oncohistones can drive cancer via diverse mechanisms
- Proteomic analysis is an invaluable complement to genomic analyses in cancer research

Figure 1: Schematized workflow for the “modification-centric” approach.

Depicted is an example in which a particular mutation confers a gain-of-function to a specific histone modifying enzyme, such as the NSD2 translocation in multiple myeloma. Histone PTM analysis can identify marks that are differentially enriched in two populations of cells that differ in expression of the mutation. Follow-up experiments that relate the presence of the PTM to expression of specific genes can often explain aspects of the phenotype associated with the mutation.

Figure 2: Identification of chromatin readers.

Peptide pulldown experiments coupled to downstream proteomics can identify readers of specific histone modifications. The purple circle represents a histone PTM of interest. The red protein represents a nonspecific interactor, the green protein represents a specific reader of the unmodified peptide, and the cyan peptide represents a specific reader of the modified peptide. Classically, this workflow utilizes concentration of peptide on beads, so that the ratio of the abundance of any protein across the two conditions indicates its relative affinity. However, if multiple peptide concentrations are tested, it may be possible to derive binding curves for each protein against each bait, providing a more quantitative description of reader behavior.

Figure 3: Schematized workflow for the “complex-centric” approach.

When an oncogenic fusion protein is identified, IP-MS can reveal protein-protein interaction networks associated with either protein individually, and those associated with the chimeric fusion protein. In this example, the translocation results in the fusion of Protein A to Protein B, creating protein AB. Though Protein A and Protein B do not participate in any of the same protein interaction networks, Protein AB forms a hybrid complex, bringing together a subset of interactors from both A and B. Proteins found to be associated with the AB fusion are candidates for mediating the oncogenic effects of fusion proteins.

Figure 4: Oncohistones and associated tumors.

Missense histone mutations have been detected in a variety of human tumors.

Table 1:
Histone Marks, Writers, and Small Molecule Inhibitors.

Inhibitors are available for a subset of histone PTM writers. Those marked with asterisks are used clinically.

PTM	Druggable Writer	Inhibitor	Reference
H3K4me1	SETD7	(R)-PFI-2	[42]
H3K4me1/2/3	MLL1	MM-401	[43]
H3K9me1/2/3	GLP/G9a	UNC0638, UNC0642	[44]
H3K27me1/2/3	EZH1/2	Tazemetostat* GSK926, GSK343 UNC1999	[45,46] [47] [48]
H3K36me3	SETD2	N-propyl sinefungin	[49]
H3K79me1/2/3	DOT1L	Pinometostat* EPZ004777, SGC0946	[50]
H4K20me1	SETD8	UNC0379	[51]
H4K20me2/3	SUV420H1/H2	A-196	[52]

Table 2:
Diverse cancers bearing translocations in chromatin modifiers.

Depicted is a selection of cancers known to bear chromosomal rearrangements affecting one or more chromatin modifiers. These are candidates for hybrid complex formation.

Fusion	Function of N-terminal partner	Function of C-terminal partner	Associated Cancer	References
BRD3/4-NUT	Histone acetylation reader	Unknown, nuclear localization	NUT midline carcinoma	[119,120]
EP400-PHF1	NuA4 histone acetyltransferase complex subunit	Polycomb complex subunit, histone methylation reader	Ossifying fibromyxoid tumor	[121]
JAZF1-SUZ12	Zinc finger DNA binding	Polycomb complex subunit	Endometrial stromal sarcoma	[122]
YWHAE-NUTM2B/E	14-3-3 family phosphorylation reader	Unknown, nuclear localization	Kidney clear cell sarcoma	[123,124]
ZMYND11-MBTD1	Histone methylation reader	Polycomb complex subunit	Acute myeloid leukemia	[125]