Cancer Epigenetics: From Mechanism to Therapy

Cell

Leading Edge
Review

Cancer Epigenetics: From Mechanism to Therapy

Mark A. Dawson[1,2] **and Tony Kouzarides**[1,*]
[1]Gurdon Institute and Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK
[2]Department of Haematology, Cambridge Institute for Medical Research and Addenbrooke's Hospital, University of Cambridge, Hills Road, Cambridge CB2 0XY, UK
*Correspondence: t.kouzarides@gurdon.cam.ac.uk
http://dx.doi.org/10.1016/j.cell.2012.06.013

The epigenetic regulation of DNA-templated processes has been intensely studied over the last 15 years. DNA methylation, histone modification, nucleosome remodeling, and RNA-mediated targeting regulate many biological processes that are fundamental to the genesis of cancer. Here, we present the basic principles behind these epigenetic pathways and highlight the evidence suggesting that their misregulation can culminate in cancer. This information, along with the promising clinical and preclinical results seen with epigenetic drugs against chromatin regulators, signifies that it is time to embrace the central role of epigenetics in cancer.

Chromatin is the macromolecular complex of DNA and histone proteins, which provides the scaffold for the packaging of our entire genome. It contains the heritable material of eukaryotic cells. The basic functional unit of chromatin is the nucleosome. It contains 147 base pairs of DNA, which is wrapped around a histone octamer, with two each of histones H2A, H2B, H3, and H4. In general and simple terms, chromatin can be subdivided into two major regions: (1) heterochromatin, which is highly condensed, late to replicate, and primarily contains inactive genes; and (2) euchromatin, which is relatively open and contains most of the active genes. Efforts to study the coordinated regulation of the nucleosome have demonstrated that all of its components are subject to covalent modification, which fundamentally alters the organization and function of these basic tenants of chromatin (Allis et al., 2007).

The term "epigenetics" was originally coined by Conrad Wad-

The information conveyed by epigenetic modifications plays a critical role in the regulation of all DNA-based processes, such as transcription, DNA repair, and replication. Consequently, abnormal expression patterns or genomic alterations in chromatin regulators can have profound results and can lead to the induction and maintenance of various cancers. In this Review, we highlight recent advances in our understanding of these epigenetic pathways and discuss their role in oncogenesis. We provide a comprehensive list of all the recurrent cancer mutations described thus far in epigenetic pathways regulating modifications of DNA (Figure 2), histones (Figures 3, 4, and 5), and chromatin remodeling (Figure 6). Where relevant, we will also emphasize existing and emerging drug therapies aimed at targeting epigenetic regulators (Figure 1).

Characterizing the Epigenome

Table 1. Chromatin Modifications, Readers, and Their Function

Chromatin Modification	Nomenclature	Chromatin-Reader Motif	Attributed Function
DNA Modifications			
5-methylcytosine	5mC	MBD domain	transcription
5-hydroxymethylcytosine	5hmC	unknown	transcription
5-formylcytosine	5fC	unknown	unknown
5-carboxylcytosine	5caC	unknown	unknown
Histone Modifications			
Acetylation	K-ac	BromodomainTandem, PHD fingers	transcription, repair, replication, and condensation
Methylation (lysine)	K-me1, K-me2, K-me3	Chromodomain, Tudor domain, MBT domain, PWWP domain, PHD fingers, WD40/β propeller	transcription and repair
Methylation (arginine)	R-me1, R-me2s, R-me2a	Tudor domain	transcription
Phosphorylation (serine and threonine)	S-ph, T-ph	14-3-3, BRCT	transcription, repair, and condensation
Phosphorylation (tyrosine)	Y-ph	SH2[a]	transcription and repair
Ubiquitylation	K-ub	UIM, IUIM	transcription and repair
Sumoylation	K-su	SIM[a]	transcription and repair
ADP ribosylation	E-ar	Macro domain, PBZ domain	transcription and repair
Deimination	R→Cit	unknown	transcription and decondensation
Proline isomerisation	P-cis⇔P-trans	unknown	transcription
Crotonylation	K-cr	unknown	transcription
Propionylation	K-pr	unknown	unknown
Butyrylation	K-bu	unknown	unknown
Formylation	K-fo	unknown	unknown
Hyroxylation	Y-oh	unknown	unknown
O-GlcNAcylation (serine and threonine)	S-GlcNAc; T-GlcNAc	unknown	transcription

Modifications: me1, monomethylation; me2, dimethylation; me3, trimethylation; me2s, symmetrical dimethylation; me2a, asymmetrical dimethylation; and Cit, citrulline. Reader domains: MBD, methyl-CpG-binding domain; PHD, plant homeodomain; MBT, malignant brain tumor domain; PWWP, proline-tryptophan-tryptophan-proline domain; BRCT, BRCA1 C terminus domain; UIM, ubiquitin interaction motif; IUIM, inverted ubiquitin interaction motif; SIM, sumo interaction motif; and PBZ, poly ADP-ribose binding zinc finger.

[a]These are established binding modules for the posttranslational modification; however, binding to modified histones has not been firmly established.

Recent improvements in the sensitivity and accuracy of mass spectrometry (MS) instruments have driven many of these discoveries (Stunnenberg and Vermeulen, 2011). Moreover, although MS is inherently not quantitative, recent advances in labeling methodologies, such as stable isotope labeling by amino acids in cell culture (SILAC), isobaric tags for relative and absolute quantification (iTRAQ), and isotope-coded affinity tag (ICAT), have allowed a greater ability to provide quantitative measurements (Stunnenberg and Vermeulen, 2011).

These quantitative methods have generated "protein recruitment maps" for histone and DNA modifications, which contain proteins that recognize chromatin modifications (Bartke et al., 2010; Vermeulen et al., 2010). Many of these chromatin readers have more than one reading motif, so it is important to understand how they recognize several modifications either simultaneously or sequentially. The concept of multivalent engagement by chromatin-binding modules has recently been explored by using either modified histone peptides (Vermeulen et al., 2010) or in-vitro-assembled and -modified nucleosomes (Bartke et al., 2010; Ruthenburg et al., 2011). The latter approach in particular has uncovered some of the rules governing the recruitment of protein complexes to methylated DNA and modified histones in a nucleosomal context. The next step in our understanding will require a high-resolution in vivo genomic approach to detail the dynamic events on any given nucleosome during the course of gene expression.

Epigenetics and the Cancer Connection

The earliest indications of an epigenetic link to cancer were derived from gene expression and DNA methylation studies. These studies are too numerous to comprehensively detail in this review; however, the reader is referred to an excellent review detailing the history of cancer epigenetics (Feinberg and Tycko, 2004). Although many of these initial studies were purely correlative, they did highlight a potential connection between epigenetic pathways and cancer. These early observations have been significantly strengthened by recent results from the International Cancer Genome Consortium (ICGC). Whole-genome

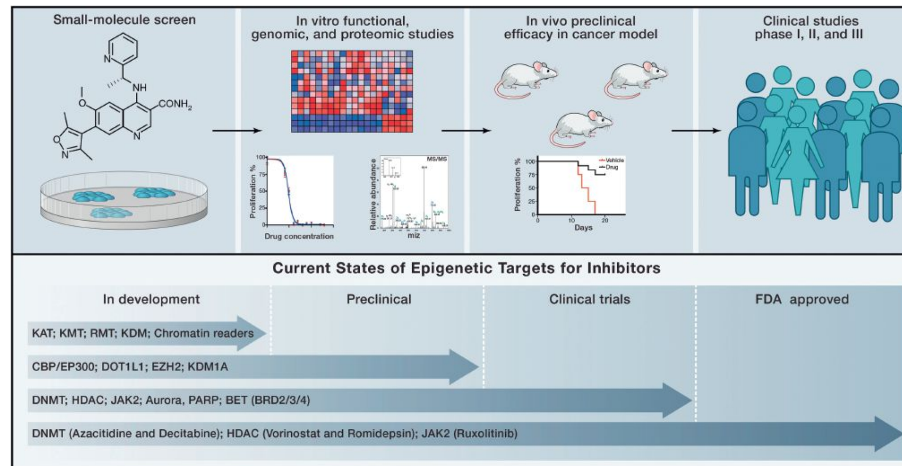


Figure 1. Epigenetic Inhibitors as Cancer Therapies

This schematic depicts the process for epigenetic drug development and the current status of various epigenetic therapies. Candidate small molecules are first tested in vitro in malignant cell lines for specificity and phenotypic response. These may, in the first instance, assess the inhibition of proliferation, induction of apoptosis, or cell-cycle arrest. These phenotypic assays are often coupled to genomic and proteomic methods to identify potential molecular mechanisms for the observed response. Inhibitors that demonstrate potential in vitro are then tested in vivo in animal models of cancer to ascertain whether they may provide therapeutic benefit in terms of survival. Animal studies also provide valuable information regarding the toxicity and pharmacokinetic properties of the drug. Based on these preclinical studies, candidate molecules may be taken forward into the clinical setting. When new drugs prove beneficial in well-conducted clinical trials, they are approved for routine clinical use by regulatory authorities such as the FDA. KAT, histone lysine acetyltransferase; KMT, histone lysine methyltransferase; RMT, histone arginine methyltransferase; and PARP, poly ADP ribose polymerase.

sequencing in a vast array of cancers has provided a catalog of recurrent somatic mutations in numerous epigenetic regulators (Forbes et al., 2011; Stratton et al., 2009). A central tenet in analyzing these cancer genomes is the identification of ‘‘driver’’ mutations (causally implicated in the process of oncogenesis). A key feature of driver mutations is that they are recurrently found in a variety of cancers, and/or they are often present at a high prevalence in a specific tumor type. We will mostly concentrate our discussions on suspected or proven driver mutations in epigenetic regulators.

For instance, malignancies such as follicular lymphoma contain recurrent mutations of the histone methyltransferase *MLL2* in close to 90% of cases (Morin et al., 2011). Similarly, *UTX*, a histone demethylase, is mutated in up to 12 histologically distinct cancers (van Haaften et al., 2009). Compilation of the epigenetic regulators mutated in cancer highlights histone acetylation and methylation as the most widely affected epigenetic pathways (Figures 3 and 4). These and other pathways that are affected to a lesser extent will be described in the following sections.

Deep sequencing technologies aimed at mapping chromatin modifications have also begun to shed some light on the origins of epigenetic abnormalities in cancer. Cross-referencing of DNA methylation profiles in human cancers with ChIP-Seq data for histone modifications and the binding of chromatin regulators have raised intriguing correlations between cancer-associated DNA hypermethylation and genes marked with ‘‘bivalent’’ histone modifications in multipotent cells (Easwaran et al., 2012; Ohm et al., 2007). These bivalent genes are marked by active (H3K4me3) and repressive (H3K27me3) histone modifications (Bernstein et al., 2006) and appear to identify transcriptionally poised genes that are integral to development and lineage commitment. Interestingly, many of these genes are targeted for DNA methylation in cancer. Equally intriguing are recent comparisons between malignant and normal tissues from the same individuals. These data demonstrate broad domains within the malignant cells that contain significant alterations in DNA methylation. These regions appear to correlate with late-replicating regions of the genome associated with the nuclear lamina (Berman et al., 2012). Although there remains little mechanistic insight into how and why these regions of the genome are vulnerable to epigenetic alterations in cancer, these studies highlight the means by which global sequencing platforms have started to uncover avenues for further investigation.

Genetic lesions in chromatin modifiers and global alterations in the epigenetic landscape not only imply a causative role for these proteins in cancer but also provide potential targets for therapeutic intervention. A number of small-molecule inhibitors have already been developed against chromatin regulators (Figure 1). These are at various stages of development, and three

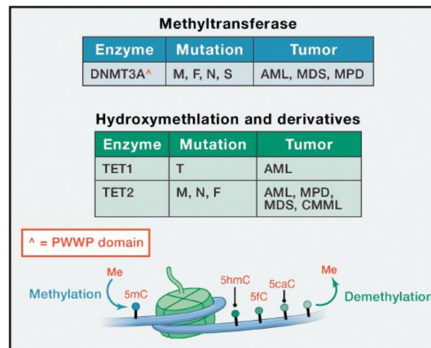
of these (targeting DNMTs, HDACs, and JAK2) have already been granted approval by the US Food and Drug Administration (FDA). This success may suggest that the interest in epigenetic pathways as targets for drug discovery had been high over the past decade. However, the reality is that the field of drug discovery had been somewhat held back due to concerns over the pleiotropic effects of both the drugs and their targets. Indeed, some of the approved drugs (against HDACs) have little enzyme specificity, and their mechanism of action remains contentious (Minucci and Pelicci, 2006).

The belief and investment in epigenetic cancer therapies may now gain momentum and reach a new level of support following the recent preclinical success of inhibitors against BRD4, an acetyl-lysine chromatin-binding protein (Dawson et al., 2011; Delmore et al., 2011; Filippakopoulos et al., 2010; Mertz et al., 2011; Zuber et al., 2011). The molecular mechanisms governing these impressive preclinical results have also been largely uncovered and are discussed below. This process is pivotal for the successful progression of these inhibitors into the clinic. These results, along with the growing list of genetic lesions in epigenetic regulators, highlight the fact that we have now entered an era of epigenetic cancer therapies.

Epigenetic Pathways Connected to Cancer

DNA Methylation

The methylation of the 5-carbon on cytosine residues (5mC) in CpG dinucleotides was the first described covalent modification of DNA and is perhaps the most extensively characterized modification of chromatin. DNA methylation is primarily noted within centromeres, telomeres, inactive X-chromosomes, and repeat sequences (Baylin and Jones, 2011; Robertson, 2005). Although global hypomethylation is commonly observed in malignant cells, the best-studied epigenetic alterations in cancer are the methylation changes that occur within CpG islands, which are present in ~70% of all mammalian promoters. CpG island methylation plays an important role in transcriptional regulation, and it is commonly altered during malignant transformation (Baylin and Jones, 2011; Robertson, 2005). NGS platforms have now provided genome-wide maps of CpG methylation. These have confirmed that between 5%–10% of normally unmethylated CpG promoter islands become abnormally methylated in various cancer genomes. They also demonstrate that CpG hypermethylation of promoters not only affects the expression of protein coding genes but also the expression of various noncoding RNAs, some of which have a role in malignant transformation (Baylin and Jones, 2011). Importantly, these genome-wide DNA methylome studies have also uncovered intriguing alterations in DNA methylation within gene bodies and at CpG "shores," which are conserved sequences upstream and downstream of CpG islands. The functional relevance of these regional alterations in methylation are yet to be fully deciphered, but it is interesting to note that they have challenged the general dogma that DNA methylation invariably equates with transcriptional silencing. In fact, these studies have established that many actively transcribed genes have high levels of DNA methylation within the gene body, suggesting that the context and spatial distribution of DNA methylation is vital in transcriptional regulation (Baylin and Jones, 2011).



Methyltransferase

Enzyme	Mutation	Tumor
DNMT3A^	M, F, N, S	AML, MDS, MPD

Hydroxymethlation and derivatives

Enzyme	Mutation	Tumor
TET1	T	AML
TET2	M, N, F	AML, MPD, MDS, CMML

Figure 2. Cancer Mutations Affecting Epigenetic Regulators of DNA Methylation

The 5-carbon of cytosine nucleotides are methylated (5mC) by a family of DNMTs. One of these, DNMT3A, is mutated in AML, myeloproliferative diseases (MPD), and myelodysplastic syndromes (MDS). In addition to its catalytic activity, DNMT3A has a chromatin-reader motif, the PWWP domain, which may aid in localizing this enzyme to chromatin. Somatically acquired mutations in cancer may also affect this domain. The TET family of DNA hydroxylases metabolizes 5mC into several oxidative intermediates, including 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC). These intermediates are likely involved in the process of active DNA demethylation. Two of the three TET family members are mutated in cancers, including AML, MPD, MDS, and CMML. Mutation types are as follows: M, missense; F, frameshift; N, nonsense; S, splice site mutation; and T, translocation.

Three active DNA methyltransferases (DNMTs) have been identified in higher eukaryotes. DNMT1 is a maintenance methyltransferase that recognizes hemimethylated DNA generated during DNA replication and then methylates newly synthesized CpG dinucleotides, whose partners on the parental strand are already methylated (Li et al., 1992). Conversely, DNMT3a and DNMT3b, although also capable of methylating hemimethylated DNA, function primarily as de novo methyltransferases to establish DNA methylation during embryogenesis (Okano et al., 1999). DNA methylation provides a platform for several methyl-binding proteins. These include MBD1, MBD2, MBD3, and MeCP2. These in turn function to recruit histone-modifying enzymes to coordinate the chromatin-templated processes (Klose and Bird, 2006).

Although mutations in DNA methyltransferases and MBD proteins have long been known to contribute to developmental abnormalities (Robertson, 2005), we have only recently become aware of somatic mutations of these key genes in human malignancies (Figure 2). Recent sequencing of cancer genomes has identified recurrent mutations in *DNMT3A* in up to 25% of patients with acute myeloid leukemia (AML) (Ley et al., 2010). Importantly, these mutations are invariably heterozygous and are predicted to disrupt the catalytic activity of the enzyme. Moreover, their presence appears to impact prognosis (Patel et al., 2012). However, at present, the mechanisms by which

enzymes. In addition to their catalytic function, many chromatin modifiers also possess "reader" domains allowing them to bind to specific regions of the genome and respond to information conveyed by upstream signaling cascades. This is important, as it provides two avenues for therapeutically targeting these epigenetic regulators. The residues that line the binding pocket of reader domains can dictate a particular preference for specific modification states, whereas residues outside the binding pocket contribute to determining the histone sequence specificity. This combination allows similar reader domains to dock at different modified residues or at the same amino acid displaying different modification states. For example, some methyl-lysine readers engage most efficiently with di/tri-methylated lysine (Kme2/3), whereas others prefer mono- or unmethylated lysines. Alternatively, when the same lysines are now acetylated, they bind to proteins containing bromodomains (Taverna et al., 2007). The main modification binding pockets contained within chromatin-associated proteins is summarized in Table 1.

Many of the proteins that modify or bind these histone modifications are misregulated in cancer, and in the ensuing sections, we will discuss the most extensively studied histone modifications in relation to oncogenesis and novel therapeutics.

Histone Acetylation. The N^{ε}-acetylation of lysine residues is a major histone modification involved in transcription, chromatin structure, and DNA repair. Acetylation neutralizes lysine's positive charge and may consequently weaken the electrostatic interaction between histones and negatively charged DNA. For this reason, histone acetylation is often associated with a more "open" chromatin conformation. Consistent with this, ChIP-Seq analyses have confirmed the distribution of histone acetylation at promoters and enhancers and, in some cases, throughout the transcribed region of active genes (Heintzman et al., 2007; Wang et al., 2008). Importantly, lysine acetylation also serves as the nidus for the binding of various proteins with bromodomains and tandem plant homeodomain (PHD) fingers, which recognize this modification (Taverna et al., 2007).

Acetylation is highly dynamic and is regulated by the competing activities of two enzymatic families, the histone lysine acetyltransferases (KATs) and the histone deacetylases (HDACs). There are two major classes of KATs: (1) type-B, which are predominantly cytoplasmic and modify free histones, and (2) type-A, which are primarily nuclear and can be broadly classified into the GNAT, MYST, and CBP/p300 families.

KATs were the first enzymes shown to modify histones. The importance of these findings to cancer was immediately apparent, as one of these enzymes, CBP, was identified by its ability to bind the transforming portion of the viral oncoprotein E1A (Bannister and Kouzarides, 1996). It is now clear that many, if not most, of the KATs have been implicated in neoplastic transformation, and a number of viral oncoproteins are known to associate with them. There are numerous examples of recurrent chromosomal translocations (e.g., *MLL-CBP* [Wang et al., 2005] and *MOZ-TIF2* [Huntly et al., 2004]) or coding mutations (e.g., *p300/CBP* [Iyer et al., 2004; Pasqualucci et al., 2011]) involving various KATs in a broad range of solid and hematological malignancies (Figure 3). Furthermore, altered expression levels of several of the KATs have also been noted in a range of cancers (Avvakumov and Côté, 2007; Iyer et al., 2004). In some cases, such as the leukemia-associated fusion gene *MOZ-TIF2*, we know a great deal about the cellular consequences of this translocation involving a MYST family member. MOZ-TIF2 is sufficient to recapitulate an aggressive leukemia in murine models; it can confer stem cell properties and reactivate a self-renewal program when introduced into committed hematopoietic progenitors, and much of this oncogenic potential is dependent on its inherent and recruited KAT activity as well as its ability to bind to nucleosomes (Deguchi et al., 2003; Huntly et al., 2004).

Despite these insights, the great conundrum with regards to unraveling the molecular mechanisms by which histone acetyltransferases contribute to malignant transformation has been dissecting the contribution of altered patterns in acetylation on histone and nonhistone proteins. Although it is clear that global histone acetylation patterns are perturbed in cancers (Fraga

Acetyltransferases

Enzyme	Mutation	Tumor
KAT3A (CBP)*	T, N, F, M	AML, ALL, DLBCL, B-NHL, TCC
KAT3B (p300)*	T, N, F, M	AML, ALL, DLBCL, TCC, Colorectal, Breast, Pancreatic
KAT6A (MOZ)+	T	AML, MDS
KAT6B (MORF)+	T	AML, Uterine leiomyoma

Readers

Reader	Mutation	Tumor
BRD1**	T	ALL
BRD3*	T	Midline carcinoma
BRD4*	T	Midline carcinoma
TRIM33*+	T	Papillary thyroid
PBRM1*	N, F, M, S, D	Renal, Breast

Figure 3. Cancer Mutations Affecting Epigenetic Regulators Involved in Histone Acetylation

These tables provide somatic cancer-associated mutations identified in histone acetyltransferases and proteins that contain bromodomains (which recognize and bind acetylated histones). Several histone acetyltransferases possess chromatin-reader motifs and, thus, mutations in the proteins may alter both their catalytic activities as well as the ability of these proteins to scaffold multiprotein complexes to chromatin. Interestingly, sequencing of cancer genomes to date has not identified any recurrent somatic mutations in histone deacetylase enzymes. Abbreviations for the cancers are as follows: AML, acute myeloid leukemia; ALL, acute lymphoid leukemia; B-NHL, B-cell non-Hodgkin's lymphoma; DLBCL, diffuse large B-cell lymphoma; and TCC, transitional cell carcinoma of the urinary bladder. Mutation types are as follows: M, missense; F, frameshift; N, nonsense; S, splice site mutation; T, translocation; and D, deletion.

Histone Methylation. Histones are methylated on the side chains of arginine, lysine, and histidine residues. Methylation, unlike acetylation and phosphorylation, does not alter the overall charge of the molecule. Lysines may be mono-, di-, or tri-methylated, and arginine residues may be symmetrically or asymmetrically methylated. The best-characterized sites of histone methylation are those that occur on lysine residues and, therefore, these will be the focus of this section. Although many lysine residues on the various histones are methylated, the best studied are H3K4, H3K9, H3K27, H3K36, H3K79, and H4K20. Some of these (H3K4, H3K36, and H3K79) are often associated with active genes in euchromatin, whereas others (H3K9, H3K27, and H4K20) are associated with heterochromatic regions of the genome (Barski et al., 2007). Different methylation states on the same residue can also localize differently. For instance, H3K4me2/3 usually spans the transcriptional start site (TSS) of active genes (Barski et al., 2007), whereas H3K4me1 is a modification associated with active enhancers (Heintzman et al., 2009). Similarly, whereas monomethylation of H3K9 may be seen at active genes, trimethylation of H3K9 is associated with gene repression (Barski et al., 2007).

The enzymatic protagonists for lysine methylation contain a conserved SET domain, which possesses methyltransferase activity. The only exception to this is hDOT1L, the enzyme that methylates H3K79. In contrast to the KATs, the histone lysine methyltransferases (KMT) tend to be highly specific enzymes that specifically target certain lysine residues. Cytogenetic studies, as well as NGS of various cancer genomes, have demonstrated recurrent translocations and/or coding mutations in a large number of KMT, including *MMSET*, *EZH2*, and *MLL* family members (Figure 4).

Whereas the oncogenic effects exerted by the MLL fusions have been extensively studied and reviewed (Krivtsov and Armstrong, 2007), an emerging area of interest is the dichotomous role of EZH2 in human malignancies. EZH2 is the catalytic component of the PRC2 complex, which is primarily responsible for the methylation of H3K27. Early gene-expression studies implicated the overexpression of EZH2 as a progressive event that conferred a poor prognosis in prostate and breast cancer (Margueron and Reinberg, 2011). These initial studies suggested that *EZH2* was an oncogene. However, NGS and targeted resequencing of cancer genomes have recently identified coding mutations within EZH2 in various lymphoid and myeloid neoplasms that have somewhat muddied the waters by suggesting both oncogenic and tumor-suppressive roles for EZH2. Heterozygous missense mutations resulting in the substitution of tyrosine 641 (Y641) within the SET domain of EZH2 were noted in 22% of patients with diffuse large B-cell lymphoma (Morin et al., 2010). Functional characterization of this mutation demonstrated that it conferred increased catalytic activity and a preference for converting H3K27me1 to H3K27me2/3, again supporting the contention that *EZH2* is an oncogene (Sneeringer et al., 2010). In contrast, loss-of-function mutations in *EZH2* gene, conferring a poor prognosis, have been described in the myeloid malignancies (Ernst et al., 2010; Nikoloski et al., 2010) and T-ALL (Ntziachristos et al., 2012; Zhang et al., 2012), suggesting a tumor-suppressive role for EZH2 in these cell lineages.

Methyltransferases

Enzyme	Mutation	Tumor
KMT2A (MLL1*+)	T, PTD	AML, ALL, TCC
KMT2B (MLL2*)	N, F, M	Medulloblastoma, Renal, DLBCL, FL
KMT2C (MLL3*)	N	Medulloblastoma, TCC, Breast
KMT3A (SETD2)	N, F, S, M	Renal, Breast
KMT3B (NSD1+^)	T	AML
NSD2+^	T	Multiple myeloma
NSD3^	T	AML
KMT6 (EZH2)	M	DLBCL, MPD, MDS

Readers

Reader	Mutation	Tumor
TRIM33+*	T	Papillary thyroid
ING1+	M, D	Melanoma, Breast
ING4+	D	HNSCC
MSH6^	M, N, F, S	Colorectal

Demethylases

Enzyme	Mutation	Tumor
KDM5A (JARID1A)+	T	AML
KDM5C (JARID1C)+	N, F, S	Renal
KDM6A (UTX)	D, N, F, S	AML, TCC, Renal, Oesophageal, Multiple myeloma

Figure 4. Cancer Mutations Affecting Epigenetic Regulators Involved in Histone Methylation
Recurrent mutations in histone methyltransferases, demethylases, and methyllysine binders have been identified in a large number of cancers. These mutations may significantly alter the catalytic activity of the methyltransferases or demethylases. In addition, as many of these enzymes also contain chromatin-reader motifs, they may also affect the ability of these proteins to survey and bind epigenetic modifications. Abbreviations for the cancers are as follows: AML, acute myeloid leukemia; ALL, acute lymphoid leukemia; B-NHL, B-cell non-Hodgkin's lymphoma; DLBCL, diffuse large B-cell lymphoma; HNSCC, head and neck squamous cell carcinoma; FL, follicular lymphoma; MDS, myelodysplastic syndromes; MPD, myeloproliferative diseases; and TCC, transitional cell carcinoma of the urinary bladder. Mutation types are as follows: M, missense; F, frameshift; N, nonsense; S, splice site mutation; T, translocation; D, deletion; and PTD, partial tandem duplication.

The precise mechanisms by which gain and loss of EZH2 activity culminate in cancers are an area of active investigation. In light of the varied roles that polycomb proteins play in

self-renewal and differentiation (Margueron and Reinberg, 2011), solution of this problem will necessitate vigilance and appreciation of the cellular context within which the mutations arise. The increased awareness of the involvement of KMTs in cancer has heightened efforts to identify specific inhibitors. These efforts will only be encouraged by the recent demonstration that small-molecule inhibition of DOT1L shows preclinical promise as a targeted therapy in MLL leukemia (Daigle et al., 2011), a disease in which aberrant DOT1L activity is ill defined but clearly involved (Krivtsov and Armstrong, 2007).

Histone Demethylation. The initial notion that histone lysine methylation was a highly stable, nondynamic modification has now been irrefutably overturned by the identification of two classes of lysine demethylases (Mosammaparast and Shi, 2010). The prima facie example, LSD1 (KDM1A), belongs to the first class of demethylases that demethylates lysines via an amine oxidation reaction with flavin adenine dinucleotide (FAD) as a cofactor. As this family of enzymes requires a protonated nitrogen to initiate demethylation, they are limited to demethylating mono- and dimethyllysine. The second and more expansive class of enzymes is broadly referred to as the Jumonji demethylases. They have a conserved JmjC domain, which functions via an oxidative mechanism and radical attack (involving Fe(II) and α-ketoglutarate). The Jumonji family does not require a free electron pair on the nitrogen atom to initiate catalysis and, therefore, unlike LSD1, they can demethylate all three methyl lysine states. Unsurprisingly, the multisubunit complexes within which these enzymes reside confer much of their target specificity. As an example, LSD1 can function as a transcriptional repressor by demethylating H3K4me1/2 as part of the corepressor for RE1-silencing transcription factor (Co-REST) complex, but its activity is linked to gene activation when it associates with the androgen receptor to demethylate H3K9me2 (Mosammaparast and Shi, 2010). Thus far, recurrent coding mutations have been noted in *KDM5A* (*JARID1A*), *KDM5C* (*JARID1C*), and *KDM6A* (*UTX*) (Figure 4). Mutations in *UTX*, in particular, are prevalent in a large number of solid and hematological cancers. Small-molecule inhibitors of the two families of histone demethylases are at various stages of development, and this interest will be spurred on by emerging preclinical data showing the therapeutic potential of compounds that inhibit LSD1/KDM1A in AML (Barretina et al., 2012; Schenk et al., 2012).

Interestingly, recent findings related to recurrent mutations in the genes encoding the metabolic enzymes isocitrate dehydrogenase-1 (*IDH1*) and *IDH2* have broad implications for the Jumonji class of demethylases, which use α-ketoglutarate (α-KG). IDH1/2 are nicotinamide adenine dinucleotide phosphate (NADP)-dependent enzymes that normally catalyze the oxidative decarboxylation of isocitrate to α-KG, which is associated with the production of NADPH. Mutations in *IDH1* and *IDH2* are seen in up to 70% of patients with secondary glioblastoma mutiforme and are also noted as recurrent mutations in a range of myeloid malignancies, most notably AML (Cimmino et al., 2011). These mutations manifest in a neomorphic enzymatic activity that results in the NADPH-dependent reduction of α-KG to 2-hydroxyglutarate (2-HG). Consequently, malignant cells with *IDH1/2* mutations may harbor 2-HG levels that are up to 100-fold higher than normal (Cimmino et al., 2011). 2-HG is a competitive inhibitor of the α-KG-dependent dioxygenases; in fact, 2-HG has been shown to adopt a near-identical orientation within the catalytic core of the JmjC domain (Xu et al., 2011). As 2-HG levels accumulate within the malignant cells, there is a purported blanket inhibition of the Jumonji class of histone demethylases. Accordingly, there is a discernable increase in histone methylation levels (Xu et al., 2011). These remarkable findings are yet to be fully investigated, and it will be important to determine whether all the Jumonji family members are equally susceptible to 2-HG inhibition. A similar question can be posed for the TET family of enzymes (see above), which also use α-ketoglutarate.

Histone Methylation Readers. The various states of lysine methylation result in considerable physicochemical diversity of lysine; these modification states are read and interpreted by proteins containing different specialized recognition motifs. Broadly speaking, the aromatic cages that engage methyllysine can be divided into two major families, the Royal Family (Tudor domains, Chromo domains, and malignant brain tumor [MBT] domains) and PHD fingers. The structural composition of these domains that allows for this diversity has recently been expertly reviewed (Taverna et al., 2007).

Analogous to the situation with bromodomain proteins, several methyllysine readers have also been implicated in cancer (Figure 4). For instance, all three isoforms of the chromodomain protein HP1 have altered expression in numerous cancers (Dialynas et al., 2008). However, thus far, no cancer-specific somatic mutations have been identified in HP1. In contrast, ING family members have had coding mutations identified in malignancies such as melanoma and breast cancer, including those that specifically target the PHD finger, which recognizes H3K4me3 (Coles and Jones, 2009). Despite these findings, neither of the aforementioned examples establishes a causal relationship between cancer and the abrogation of methyllysine binding at chromatin. The best example of this, and indeed a proof of principle for therapeutically targeting methyllysine binders, has recently been shown in a specific form of AML (Wang et al., 2009). Leukemia, induced by the fusion of NUP98 with the PHD finger containing part of JARID1A or PHF23, can be abrogated by mutations that negate the ability of the PHD finger to bind H3K4me3. Functional compensation of this effect can be provided by other PHD fingers that recognize this modification, but not those that do not bind H3K4me3. Moreover, mechanistic insights were provided, demonstrating that chromatin binding of the fusion protein inhibits the deposition of H3K27me3, which leads to the continued expression of critical hematopoietic oncogenes such as *HoxA9*, *Meis1*, and *Pbx1* (Wang et al., 2009). In light of these findings, and as result of the structural diversity present in methyllysine-binding modules, it is likely that small molecules that disrupt this important protein-protein interaction may be effective anticancer agents.

Histone Phosphorylation. The phosphorylation of serine, threonine, and tyrosine residues has been documented on all core and most variant histones. Phosphorylation alters the charge of the protein, affecting its ionic properties and influencing the overall structure and function of the local chromatin environment. The phosphorylation of histones is integral to essential cellular processes such as mitosis, apoptosis, DNA repair,

mutated histone proteins get incorporated into nucleosomes? How do these mutated proteins influence the function of histone chaperones, nucleosome assembly, stability, and mobility? One possibility uncovered from these studies suggests that telomere maintenance and heterochromatin stability may be compromised as a consequence of the H3.3 mutations. Several of these pediatric glioblastoma multiforme (GBMs) also harbored mutations in the ATRX/DAXX chromatin-remodeling complex, which is responsible for the deposition of H3.3. These tumors with mutations in *H3F3A/ATRX/DAXX* were associated with increased alternative lengthening of telomeres and genomic instability (Schwartzentruber et al., 2012). The *ATRX/DAXX* mutations described here are also a seminal feature of pancreatic neuroendocrine tumors (Jiao et al., 2011) and highlight emerging evidence suggesting that mutations in members of chromatin-remodeling complexes are a common feature in human malignancy.

Chromatin Remodelers

The myriad of covalent modifications on the nucleosome often provides the scaffold and context for dynamic ATP-dependent chromatin remodeling. Based on their biochemical activity and subunit composition, the mammalian chromatin-remodeling complexes can be broadly split into four major families: the switching defective/sucrose nonfermenting (SWI/SNF) family, the imitation SWI (ISWI) family, the nucleosome remodeling and deacetylation (NuRD)/Mi-2/chromodomain helicase DNA-binding (CHD) family, and the inositol requiring 80 (INO80) family. These enzymes are evolutionarily conserved and use ATP as an energy source to mobilize, evict, and exchange histones. Each of these families has distinct domain structures and is populated by members that contain various chromatin reader motifs (SANT domains, bromodomains, and chromodomains) that confer some regional and context specificity to their chromatin-remodeling activities (Wang et al., 2007).

Several members from the various chromatin-remodeling families, such as SNF5 (Versteege et al., 1998), BRG1 (Wilson and Roberts, 2011), and MTA1 (Li et al., 2012), were known to be mutated in malignancies, raising the possibility that they may be bone fide tumor suppressors (Figure 6). Strong evidence in support of this contention has now emerged from the sequencing of cancer genomes. These efforts have highlighted high-frequency mutations in several SWI/SNF complex members in a range of hematological (Chapman et al., 2011; Morin et al., 2011) and solid malignancies (Gui et al., 2011; Jones et al., 2010; Tan et al., 2011; Varela et al., 2011; Wang et al., 2011). The prevalence of these mutations would suggest that many of the members of these complexes are involved in the development and maintenance of cancer; however, functional insights into the mechanisms of oncogenesis are only just beginning to emerge. It is clear that the SWI/SNF complexes have several lineage-specific subunits and interact with tissue-specific transcription factors to regulate differentiation. They also have a reciprocal and antagonistic relationship with the polycomb complexes. One possibility, which remains to be formally established, is that mutations in SWI/SNF members potentiate malignancy by skewing the balance between self-renewal and differentiation. Recent data would also suggest a role for the SWI/SNF complexes in regulating cell-cycle progression, cell motility, and nuclear hormone signaling (Wilson and Roberts, 2011).

SWI/SNF

Gene	Mutation	Tumor
BRG1*	N, M, F, D	Lung, Rhabdoid, Medulloblastoma, Breast, Prostate, Pancreas, HNSCC
BRM*	N, M,F	HNSCC
ARID1A	N, F, M, T	OCC, Endometroid, Renal, Gastric, Breast, Medulloblastoma, TCC
ARID1B	F, M, D	Breast
ARID2	N, F, S	Hepatocellular carcinoma
SNF5	D, N, F, S, T	Rhabdoid, Familial Schwannomatosis, Chondrosarcoma, Epethioloid sarcoma, Meningioma, Chordoma, Undifferentiated sarcoma
PBRM1*	N, F, M, S, D	Renal, Breast
BCL7A	T, M	B-NHL, Multiple myeloma
BAF60A	M	Breast

* = Bromodomain

Chromatin remodeling

Figure 6. Cancer Mutations Affecting Members of the SWI/SNF Chromatin-Remodeling Complex

SWI/SNF is a multisubunit complex that binds chromatin and disrupts histone-DNA contacts. The SWI/SNF complex alters nucleosome positioning and structure by sliding and evicting nucleosomes to make the DNA more accessible to transcription factors and other chromatin regulators. Recurrent mutations in several members of the SWI/SNF complex have been identified in a large number of cancers. Abbreviations for the cancers are as follows: B-NHL, B-cell non-Hodgkin's lymphoma; HNSCC, head and neck squamous cell carcinoma; OCC, ovarian clear cell carcinoma; and TCC, transitional cell carcinoma of the urinary bladder. Mutation types are as follows: M, missense; F, frameshift; N, nonsense; S, splice site mutation; T, translocation; and D, deletion.

Genetic evidence from mouse models has confirmed that altered expression of these purported tumor suppressors can increase the propensity to develop cancer. In the case of *BRG1*, even haploinsufficiency results in increased tumors (Wilson and Roberts, 2011). However, despite the wealth of information implicating the SWI/SNF complexes in cancer (Figure 6), there is no mechanistic evidence to demonstrate that altered chromatin remodeling due to aberrant chromatin binding or loss of ATPase activity is involved.

Noncoding RNAs

The high-throughput genomic platforms have established that virtually the entire genome is transcribed; however, only ~2% of this is subsequently translated (Amaral et al., 2008). The remaining "noncoding" RNAs (ncRNAs) can be roughly categorized into small (under 200 nucleotides) and large ncRNAs. These RNAs are increasingly recognized to be vital for normal development and may be compromised in diseases such as cancer. The small ncRNAs include small nucleolar RNAs (snoRNAs), PIWI-interacting RNAs (piRNAs), small interfering RNAs (siRNAs), and microRNAs (miRNAs). Many of these families show a high

degree of sequence conservation across species and are involved in transcriptional and posttranscriptional gene silencing through specific base pairing with their targets. In contrast, the long ncRNAs (lncRNAs) demonstrate poor cross-species sequence conservation, and their mechanism of action in transcriptional regulation is more varied. Notably, these lncRNAs appear to have a critical function at chromatin, where they may act as molecular chaperones or scaffolds for various chromatin regulators, and their function may be subverted in cancer (Wang and Chang, 2011).

One of the best-studied lncRNAs that emerges from the mammalian HOXC cluster but invariably acts in *trans* is HOTAIR. HOTAIR provides a concurrent molecular scaffold for the targeting and coordinated action of both the PRC2 complex and the LSD1-containing CoREST/REST complex (Wang and Chang, 2011). HOTAIR is aberrantly overexpressed in advanced breast and colorectal cancer (Kogo et al., 2011; Wang and Chang, 2011), and manipulation of HOTAIR levels within malignant cells can functionally alter the invasive potential of these cancers by changing PRC2 occupancy (Wang and Chang, 2011). An equally intriguing example that has broad implications for both normal development and aberrant targeting of chromatin complexes in cancer is the lncRNA HOTTIP. In contrast to HOTAIR, HOTTIP is expressed from the mammalian *HOXA* cluster and acts in *cis* to aid in the transcriptional activation of the 5′ *HOXA* genes (Wang and Chang, 2011). HOTTIP, by means of chromatin looping, is brought into close proximity of the 5′ *HOXA* genes and recruits MLL1 complexes to lay down H3K4me3 and potentiate transcription. Given that the 5′ *HOXA* cluster plays a seminal role in development and maintenance of a large number of leukemias, these findings raise the possibility that abnormal expression and/or function of HOTTIP may be a feature of these diseases.

Discerning the molecular mechanisms and nuances of RNA-protein interactions is a pivotal area of chromatin research, as the stereochemical nature of these interactions may in the future lend itself to specific targeting by innovative small molecules as cancer therapies.

Perspective and Conclusions

Information from global proteomic and genomic techniques has confirmed many of the hypotheses regarding the molecular causes of cancer, but it has challenged others. The principal tenet in oncology—that cancer is a disease initiated and driven by genetic anomalies—remains uncontested, but it is now clear that epigenetic pathways also play a significant role in oncogenesis. One concern had been that the endpoint of these pathways may not necessarily be epigenetic. However, these concerns are ameliorated by the multiplicity of mutations in epigenetic regulators, including chromatin-remodeling complexes, and the observation that histones themselves are mutated at sites of key modifications in cancer. In fact, it is now irrefutable that many of the hallmarks of cancer, such as malignant self-renewal, differentiation blockade, evasion of cell death, and tissue invasiveness are profoundly influenced by changes in the epigenome.

Despite these assertions, there are still many questions to be answered before we can use our current basic knowledge in the clinical arena. The first important issue is that of selectivity. How can ubiquitously expressed epigenetic regulators serve as selective targets? The answer may lie in the fact that epigenetic components control a small number of genes instead of having global effects on gene expression. For example, the BET protein inhibitors alter only a few hundred genes, and these genes differ depending on cell type (Dawson et al., 2011; Nicodeme et al., 2010). Thus, these drugs can disrupt a selective set of genes. What remains uncertain and imperative to now learn is how these epigenetic regulators are targeted to these "essential" genes and what makes these genes solely reliant on certain epigenetic regulators.

Related to this issue is the observation that epigenetic inhibitors lead to dramatic effects in malignant cells, though their normal counterparts remain largely unaltered. This suggests that, during normal homeostasis, epigenetic regulators function in a multitiered and semiredundant manner, but in cancer, they may be required to maintain the expression of a few key target genes. A slight tip in the balance of this regulation is sufficient to result in a cell catastrophe. This "epigenetic vulnerability" of certain cancer cells in many ways mirrors the age old axiom of "oncogene addiction" (Weinstein, 2002). Some cancer cells are reliant on specific epigenetic pathways, whereas normal cells have alternative compensating pathways to rely on.

Finally, it is now also evident from both clinical and preclinical studies that hematopoietic malignancies are clearly more vulnerable to epigenetic interventions than solid malignancies. Thus, not all cancers are equally susceptible to epigenetic therapies. The biology underpinning this observation urgently warrants our attention if epigenetic therapies are to be more widely applicable. Broadly speaking, even aggressive hematopoietic malignancies, such as AML, appear to harbor as few as ten coding mutations; in contrast, the cancer genomes of solid malignancies appear to be vastly more complex. Furthermore, the in vivo niche occupied by hematopoietic cells offers a very different environment for drug exposure, and hematopoietic cells may metabolize these drugs differently than other tissues. Could these intrinsic cellular differences account for the varied efficacy of these agents? Are these therapies being used appropriately in the solid malignancies?

This latter question raises the more fundamental issue of rationally designed combination epigenetic therapies. It is likely that many of these new epigenetic drugs offer synergistic benefits, and these new therapies may also synergize with conventional chemotherapies. This strategy of combination therapy may not only increase therapeutic efficacy but also reduce the likelihood of drug resistance.

The plethora of genetic lesions in epigenetic regulators offers many possible targets for drug discovery and will no doubt attract the attention of the pharmaceutical industry. However, given the expense of the drug discovery process, what should guide the choice of target? The "drugability" of enzymes has traditionally biased this choice, but the current success of targeting acetyl-readers may propel other modification readers (e.g., methyl-readers) as the candidates of choice. In addition, one should not rely solely on the existence of genetic lesions to guide the target for drug discovery. There are no genetic lesions reported in histone deacetylases, yet clinically safe and effective