Origin and Mechanisms of DNA Methylation Dynamics in Cancers

Hariharan Easwaran and Stephen B. Baylin

Contents

Abstract DNA methylation is a key epigenetic mark in the heritable regulation of gene expression, with important roles in normal development and disease. Genomewide alterations in DNA methylation patterns are universal feature across cancers. Studies in the last few years have shown that similar alterations occur during various normal physiological processes, such as aging. Understanding mechanisms involved in DNA methylation alterations is critical for understanding cancer etiology. In this chapter, we discuss recent work on the nature of the genomic region-specific DNA methylation alterations, its functional implications, and the mechanisms underlying these alterations.

Keywords DNA methylation · CpG islands · Hypermethylation · Hypomethylation · Cancer · Epigenetics · Chromatin · Tumor suppressor · Aging · Inflammation

H. Easwaran (✉) · S. B. Baylin
Department of Oncology, The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, The Johns Hopkins University School of Medicine, Baltimore, MD, USA
e-mail: heaswar1@jhmi.edu; sbaylin1@jhmi.edu

S. Jurga, J. Barciszewski (eds.), *The DNA, RNA, and Histone Methylomes*, RNA Technologies, https://doi.org/10.1007/978-3-030-14792-1_2

1 Introduction

Cancer is an outcome of deviant gene expression programs, which helps hijacking every normal cellular process to the advantage of cancer cells. This relies on tilting the balance of oncogenic and tumor suppressive mechanisms by the altered, but heritable, gene expression programs. Epigenetic processes are central mechanisms involved in the maintenance of regulated gene expression programs between cell generations. Aberrant alterations to the epigenetic machinery are universally observed across cancers, and these alterations have prominent roles in tumorigenesis. Recent studies have revealed how epigenetic mechanisms may undergo perturbations as a result of genetic alterations, exposure to environmental agents and during the normal course of aging. These perturbations lead to changes in gene expression program that create the permissive state for functioning of cancer driver mutations. The sum effect of perturbations to epigenetic machinery and the occurrence of key cancer driver mutations lead to the uncontrolled growth and dissemination of cancer cells.

A vast range of epigenetic mechanisms are involved that can cause oncogenic and tumor suppressor inactivation. These range from a plethora of histone modifications and DNA modifications, the end result of which is the fine control of gene expression by modulating accessibility of protein factors to the chromatin, the higher order chromatin arrangement and nuclear architecture. Among these, DNA methylation is one of the foremost important functional epigenetic mark in normal development and cancers. Recent studies have highlighted other key derivative modifications of cytosine methylation (example 5-hydroxymethylcytosine), which may play important roles. This review will focus on DNA methylation, and chiefly discuss the generation and functional impact of aberrant DNA modifications in cancers. After a brief overview on epigenetics, this review will highlight the long-known importance of DNA methylation abnormalities in cancers. With this as the pivot, and special emphasis on DNA hypermethylation events, we will discuss its possible origins and functional implications, and the molecular machinery involved in methylation.

2 Epigenetics Overview

The current definition of epigenetics, i.e. a stable and heritable change in gene expression without any changes in DNA sequence (Bird 2007), implies that the molecular determinants of the epigenetic mark be faithfully replicated along with DNA replication at every round of replicative cell division. Chromatin modifications, such as histone post-translational modifications (PTMs) methylation, acetylation (to name a few), and DNA methylation are major constituents of epigenetic information subject to heritable silencing (Almouzni and Cedar 2016). Historically, DNA methylation is the first proposed epigenetic mark, which due to its biochemical properties—namely covalent linkage to DNA and a repertoire of enzymes that

maintain the methylation at every round of DNA replication (discussed later)—is the most agreed upon modification that fits the definition of epigenetic information. In mammals, DNA methylation mainly occurs as a covalent modification of cytosines in the context of the palindromic 5′-CpG-3′ dinucleotide (^{m}CpG), and is propagated on both parent and nascent strands after DNA replication. Further a complex relationship of DNA methylation with chromatin modifications, involving mutually exclusive and inclusive interactions (Kouzarides 2007), is proposed to enable methylation as a stable carrier of epigenetic information. For example, during development, differentiation and disease, the transcription factor (TF)-network and associated chromatin modifiers establish the epigenome defined by chromatin modifications. Once chromatin modifications and ensuing gene expression patterns are established, DNA methylation could provide a basis for an efficient way to re-establish this information during subsequent cell division cycles and mediate heritable transcriptional silencing.

3 DNA Methylation Abnormalities in Cancer

Due to the widespread gene expression changes in cancers, it is not surprising that the epigenome is highly perturbed compared to normal cells, which help maintain the altered expression state over the course of cancer initiation and progression. Thus, all forms of epigenetic information, viz. DNA methylation, histone PTMs, nucleosomal positioning and higher order chromosomal structure, are altered in cancers. Aberrant DNA methylation in cancers is an early change showing genomic region-specific patterns of gains and losses in DNA methylation. Recent advances in mapping methylation pattern across the whole genome ("methylome") have provided deep insights into its normal composition and large-scale alteration in cancers. In this section we highlight the key regions in the genome that harbor altered DNA methylation and how these alterations play important roles during tumorigenesis.

3.1 Global Hypomethylation

The earliest reported DNA methylation change in cancers is that of methylation losses at structural elements of genes. Normal human cells have about 70–80% of CpG residues in the genome methylated, while in cancers this reduces drastically (Ehrlich and Lacey 2013). It is now well established that gene bodies, inter-genic regions and various repetitive elements (like LINE-1) undergo hypomethylation in cancers (Ehrlich and Lacey 2013). In the earlier stages of tumorigenesis, hypomethylation can cause LOH by inducing genomic instability, but in later stages it suppresses tumor formation possibly by preventing epigenetic silencing of tumor suppressor genes (discussed later) (Yamada et al. 2005). Global DNA hypomethylation has been shown to disrupt various aspects of the normal regulation of the genome—activation

of oncogenes, altered transcription start sites, loss of imprinting, genomic instability by heterochromatin loss and reactivation of transposable elements (Chen et al. 1998; Eden et al. 2003; Ehrlich and Lacey 2013; Gaudet et al. 2003; Holm et al. 2005; Hur et al. 2014). Each of these abnormalities plays roles in tumorigenesis by altering gene expression, destabilizing the genome and increasing mutational rates.

Early studies analyzing total methylated CpG content in normal and cancer tissues have shown that DNA hypomethylation in cancers is prevalent equally across the repetitive and unique DNA sequence fractions of the genome. In accordance with the nature of genomic elements that are hypomethylated, genomewide DNA methylation analyses have confirmed these early observations by showing that hypomethylation occurs in about 50% of the genome in blocks of contiguous genomic regions termed partially-methylated domains (PMDs) that are greater than >100 kb in size (Berman et al. 2011; Hansen et al. 2011). The majority (70–80%) of the genomic CpG sites are otherwise methylated, thus resulting in genomic regions containing highly methylated domains (HMDs) interspersed by the PMDs. Key features of PMDs are that they have low gene density, they are embedded in the late replicating domains, and their boundaries associate with the nuclear lamina domain (LAD) and insulator proteins like CTCF (Bergman and Cedar 2013; Berman et al. 2011; Hansen et al. 2011) (Fig. 1a). PMDs are present in differentiated primary cells and immortalized cell lines, but not in embryonic stem cells (Lister et al. 2011). Recent work has determined a DNA sequence signature, individual units of the sequence "WCGW" (W is adenine or thymine), that is most prone to hypomethylation. Mapping CpG methylation changes in this signature sequence throughout the genome has revealed that PMDs are conserved and universal feature of all normal lineage committed proliferating primary tissues and cells in culture (Zhou et al. 2018). The degree of the PMD-HMD contrast, driven by the depth of hypomethylation of the signature sequence, is very pronounced in cancer cells compared to their normal cell counterparts. The contrasting retention of methylation in HMD and loss of methylation in PMD indicates that the regulators and enzyme machinery responsible for epigenetic maintenance have differential activity in different regions of the genome. Especially it indicates that the methylation aberrations in cancers are not due to a general global loss or gain in activity of the DNA methyltransferase (DNMT) enzymes responsible for catalyzing the methylation of cytosine, but that their local recruitment and activity in defined regions in the genome is altered.

As mentioned above, detailed analyses of the nature of the PMD domains with respect to other genomic features have revealed that they correlate best with late replicating domains in the genome (Fig. 1a). During S-phase when DNA is replicated, the newly replicated, unmethylated daughter strands are re-methylated to copy the methylation pattern of the parent strand by the DNMT1 enzyme (Jones and Liang 2009). Given the high rate at which DNA is replicated in the replication fork (about 0.03 s per nucleotide) (Jackson and Pombo 1998) and the slow rate of DNA methylation by recombinant DNMT1 (~73–433 s per methyl group transfer) (Pradhan et al. 1999), maintenance methylation by DNMT1 at each round of cell division needs to keep up with the quick rate of DNA synthesis. DNMT1 acts in

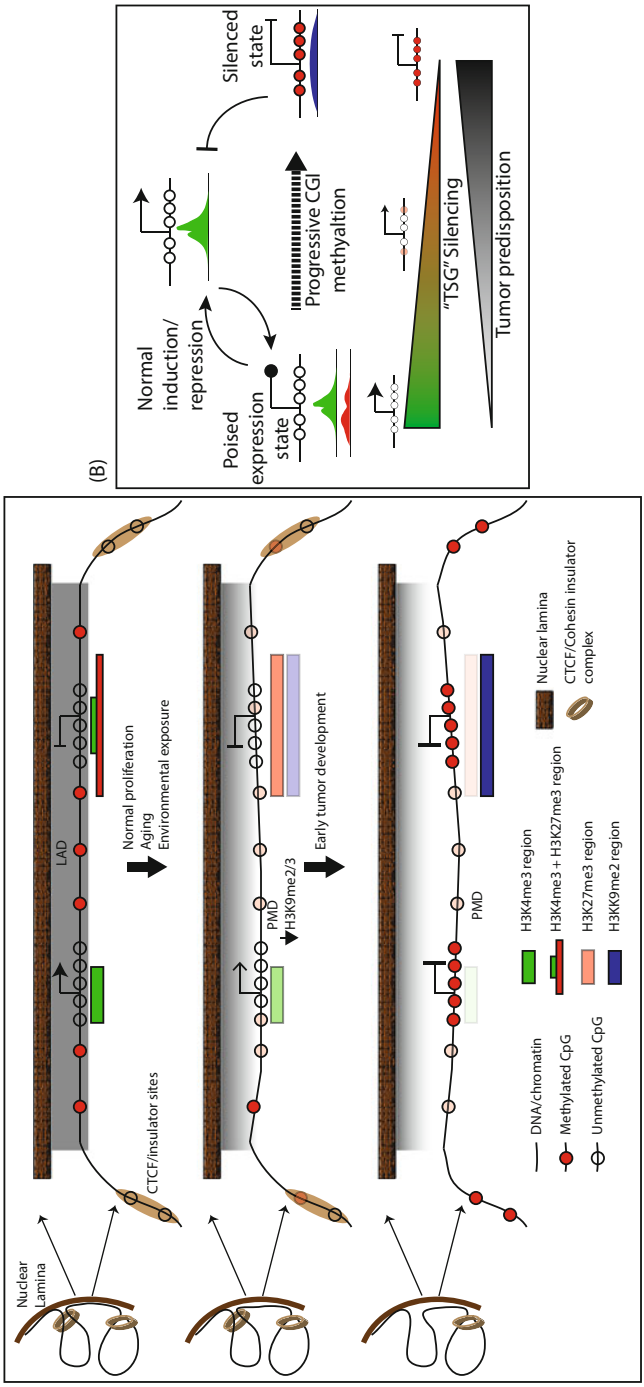

Fig. 1 (**a**) Regions of DNA showing aberrant methylation changes. DNA/chromatin is normally organized in domains associated with the nuclear lamina (lamin associated domains, LAD), and other regions forming higher ordered looped domains in the nuclear interior (left panel). During progressive cell divisions (in association with aging, environmental exposure, inflammation, etc.), DNA methylation decreases across the LAD forming the partially methylated domains (PMD). Simultaneously local promoter CGI (increased density of red circles) gain methylation, which is associated with silencing. Increasing DNA methylation and decreasing H3K4me3 and H3K27me3 at CGI promoters may be involved in a feedback regulatory mechanism causing progressive gain in DNA methylation and loss of these histone marks, causing increased gene silencing. Progressive loss of CpG methylation across the PMD, along with changes in other key heterochromatic histone modifications (like loss of H3K9me2/me3) is associated with loss of lamin association, and global gene expression changes.

concert with the replication machinery via being recruited to these sites by combined interaction with hemi-methylated DNA, PCNA and UHRF1 (Chuang et al. 1997; Easwaran et al. 2004; Leonhardt et al. 1992; Sharif et al. 2007). At these sites DNMT1 continues to remain associated with the late-replicated DNA post S-phase in the G2/M phases (Easwaran et al. 2004, 2005), and its methylation maintenance function continues even during G2 phase (Schermelleh et al. 2007). The fact that the hypomethylated PMDs are embedded in late replicating domains most likely indicates that the maintenance function is not efficient, and as a result the methylation mark is eroded at every cell cycle. PMDs, which encompasses both unique (containing genes and regulatory elements) and repetitive regions of the genome (retroviral elements), could be the central players subject to altered gene expression and genomic instability during progressive mitotic cycles in normal aging and cancer cells.

Beyond the promiscuous nature of methylation maintenance at the PMDs, whether or not other mechanisms involving the epigenetic machinery may actively regulate these regions remains to be studied. In somatic cells, maintenance of DNA methylation relies on the combined activity of the three major DNA methyltransferases, DNMT1, DNMT3A and B (Jones and Liang 2009), and the demethylation activities of TET enzymes (Wu and Zhang 2014). Whether or not activity of these enzymes, or the multitude of epigenetic regulators mutated in cancers (Dawson 2017), may accentuate hypomethylation in PMDs needs to be determined. Interestingly the PMDs are enriched for genes that undergo focal CpG island (CGI) promoter hypermethylation (discussed in next section). In the normal scenario, PMDs form during differentiation (Lister et al. 2011), but without focal CGI-hypermethylation (Berman et al. 2011). Mechanistically it is pertinent to explore if the global hypomethylation and focal hypermethylation are coupled processes occurring during aging and cancer. Further, since current studies suggest that the hypomethylation at PMDs are generated due to methylation erosion during successive cell division cycles, it is suggested that the degree of hypomethylation in PMD could indicate the mitotic history of a cell (Zhou et al. 2018). Highly proliferative tumors will thus have deeper PMD-HMD domains. The utility of using PMD as a "mitotic clock" is an exciting possibility to track history and rate of cell division in tumors at primary tumor and metastatic sites.

Fig. 1 (continued) Methylation of boundary insulator elements cause inhibition of CTCF binding resulting in further large-scale structural changes. (**b**) Promoter CGI methylation. Majority CGI promoters methylated in cancers are usually marked with H3K4me3 (green) and H3K27me3 (red) (bivalent state) in normal stem progenitor cells, representing a poised expression state. During normal cellular homeostasis (normal differentiation, exposure to stress, etc.) these genes are subject to regulated induction and repression. Progressive promoter CGI methylation accumulation causes promoter silencing and non-responsiveness to induction. Promoter silencing by CGI methylation is also associated with heterochromatic H3K9me2/3 marks (blue). A "continuum" model of "TSG" silencing wherein progressive silencing of important developmental and differentiation regulators (shown in the bottom) will lead to a gradient of expression states that will accordingly be associated with increasing predisposition to tumorigenesis

3.2 Gene Promoter CpG-Hypermethylation/Hypomethylation

About 60% of mammalian gene promoters have more than the expected occurrence of the palindromic CpG dinucleotide sequence in narrow contiguous regions around the transcription start site (TSS), termed CpG islands (CGI) (Suzuki and Bird 2008). The concept of CGI was derived from early observations that the methylation-sensitive restriction enzyme HpaII recognizing and cleaving CCGG sequence generates unexpectedly small fragments in the mouse genome (HpaII-tiny fragments, HTF) indicating that these sites are concentrated in certain regions of the genome (Illingworth and Bird 2009). These regions, termed "islands" are distinct from the rest of the genome in that they: (1) are unique sequences ranging in size ~1kb; (2) contain tenfold higher HpaII sites; (3) are GC-rich without depletion of CpG-dinucleotide sequence (which otherwise is depleted in the vertebrate genome); (4) are unmethylated in all normal tissues (except inactive-X chromosome, imprinted and germline genes); (5) are generally present in the 5′-promoter region of genes. A more formal and practical definition of CGI based on genomic sequencing is that these are regions around gene promoters devoid of Alu-repetitive elements, which are greater than 500 bp, with a GC content greater than 55% and observed CpG/expected CpG ratio above 0.65 (Takai and Jones 2002). CGI in their unmethylated state are subject to regulated gene expression, while in their methylated state are subject to permanent silencing (Baylin and Jones 2016). Thus whereas most protein-coding genes are in a transcriptionally permissive chromatin state harboring active chromatin marks (like H3K4me3 and H3K9, 14Ac) with productive transcriptional initiation by RNA Pol II, methylated genes are distinctly devoid of the active marks (Sen et al. 2016) and RNA Pol II indicating a non-permissive chromatin state (Deaton and Bird 2011).

Absence/loss of methylation of CpGs in promoters is associated with gene activation, while presence/gain of methylation is associated with gene silencing (Fig. 1b). This inverse relation is better correlated for genes with CGI (Baylin and Jones 2016). Although also observed for non-CGI genes (Han et al. 2011; Hartung et al. 2012), in somatic cells methylation of non-CGI promoters does not rule out gene expression (Weber et al. 2007). Thus it appears that the role of methylation in non-CGI promoters, vis-à-vis the role of histone modifications, is less relevant, and needs further studies. Methylation of the CpGs in the CGI and in the sequences around TSS promoter region alters chromatin structure, which inhibits binding of the transcription machinery (Baubec et al. 2013; Baylin and Jones 2016; Deaton and Bird 2011). The most well studied methylation changes in cancers are CGI-promoter alterations because ~98% of the CGI promoters are unmethylated in normal somatic cells, and because of the established inverse correlation between expression and CGI methylation. Both aberrant hypomethylation and hypermethylation of the CGI promoter genes are observed in cancers. Hypomethylation of certain CGI promoters in cancers is associated with activation of growth-promoting cell cycle regulators (Mazor et al. 2015) and genes otherwise expressed specifically in the germline, whose expression in tumors may be associated with proto-oncogenic functions (Van

Tongelen et al. 2017). On the other hand, larger numbers of genes in cancers harbor promoter CGI hypermethylation causing de novo gene silencing, or for many genes that are already silenced, CGI methylation will result in blocking of induction of these genes in response to normal differentiation cues (Fig. 1b).

The effects of CGI hypomethylation and hypermethylation at promoters in cancers parallel those of genetic alterations, namely the "gain-of-function" oncogenic activation and "loss-of-function" tumor suppressor gene (TSG) inactivation, respectively. Since CGI hypermethylation is a predominant epigenetic change in cancers, affecting promoters of about 5–10% of CGI containing genes, the mechanisms underlying their methylation and their roles in tumorigenesis are of great interest and studied extensively. In this regard, a key question is whether a given hypermethylated gene is a "driver" of the cancer phenotype. Importantly, in this regard, DNA methylation has been observed in almost every genetically identified TSG (Herman and Baylin 2003). In the context of classic TSG's, these harbor truncating genetic alterations, such as mutations, insertions and deletions, completely inactivating the genes. Since deposition of DNA methylation marks at regulatory elements mainly causes gene silencing, this constitutes a key alternative mechanism to inactivate tumor suppressor genes, wherein one copy of the TSG is inactivated by genetic alteration and the second hit is an epigenetic alteration. A strict definition of TSG requires that these genes oppose mechanisms involved in promoting tumorigenesis and that both copies of the genes should be inactivated in a "two-hit" model of tumor initiation (Knudson 2001). Since deposition of DNA methylation mark at regulatory elements mainly causes gene silencing, in many cancer cases it constitutes a key alternative mechanism to inactivate tumor suppressor genes, wherein one copy of the TSG is inactivated by genetic alteration and the second hit is an epigenetic alteration (Herman and Baylin 2003). Importantly, DNA methylation has been observed in almost every genetically identified TSG. Often important driver genes, such as *CDKN2A* and *MLH1*, are inactivated by methylation than mutation, and far more number of genes than that are genetically altered are epigenetically altered (Gao and Teschendorff 2017; Schuebel et al. 2007).

In addition to the above classic TSG's, many more promoter hypermethylated genes are seldom mutated in cancer raising the driver versus passenger question more poignantly than for the above discussed genes. Many of the affected genes identified belong to important cancer processes. Important insight for the importance of the bulk of hypermethylated genes comes from examining their potential relevance in tumor signaling pathways. In the case of genetic mutations, extensive characterization has indicated that about 2–8 mutations typically represent driver genes in any given tumor, which play key roles in tumor initiation and stepwise progression (Vogelstein et al. 2013). Importantly, more genes in key cancer related pathways are inactivated by epigenetic silencing than by mutations (Schuebel et al. 2007). How many of the methylated genes and what roles they exactly play in the context of the cancer driver events is largely unknown. A compelling case for epigenetic changes playing important roles in tumor development are pediatric cancers which have lower mutational load (Lawrence et al. 2013; McKenna and Roberts 2009), particularly the ependymomas, which are childhood brain tumors

that lack genetic alterations but show extensive DNA methylation changes of differentiation genes (Mack et al. 2014). Although these studies show the indisputable role for epigenetic changes in cancers, a lot needs to be learnt about the roles of the affected genes, especially during the course of tumor development.

To understand the function of CGI-promoter hypermethylated genes, it is important to consider that many of these genes do not fit the classic definition of TSG proposed in "Knudson's two-hit" model, in that its complete inactivation might not be the central driver of tumor development. A refined definition of TSG by Knudson and colleagues may help define the extent of the role of epigenetic silencing of such genes in tumorigenesis (Berger et al. 2011). In the revised definition, TSG is viewed as a concept wherein full inactivation of involved genes is not required, but rather a "continuum" of partial silenced states may play critical roles in tumorigenesis. This view of the role of epigenetic modulation of gene silencing is especially relevant when considering the multitude of promoter CGI hypermethylated genes without direct roles in cell cycle and DNA repair checkpoints (classic tumor suppression), but those that are involved in developmental pathways (Easwaran et al. 2012; Ohm et al. 2007; Schlesinger et al. 2007; Widschwendter et al. 2007) and other key processes such as immune checkpoint mechanisms (Wrangle et al. 2013). In this regard, it is interesting to note that 60–70% of the de novo CGI-promoter methylated genes in cancer cells are differentiation and developmental regulators marked by PcG-associated silencing in normal embryonic stem cells and adult progenitor cells (Easwaran et al. 2012). By the combined actions of trithorax (responsible for the activating H3K4me3 modification) and polycomb (PcG, responsible for the activating H3K27me3 modification) mediated bivalent chromatin marks (simultaneous presence of both H3K4me3 and H3K27me3 in the same nucleosome) in embryonic and adult stem cells, these genes are maintained at poised expression state amenable to repression or activation upon normal differentiation cues (Bernstein et al. 2006) (Fig. 1b). Thus many of the genes that are methylated in cancers are already in silent/low-expression state in the corresponding normal cells (Easwaran et al. 2012). This observation has called into question the actual benefit of promoter CGI methylation in cancer development (O'Donnell et al. 2014). However, as mentioned above, it is very important to recognize that many of the cancer-specific, promoter CGI methylated genes are regulators of development and differentiation, and expressed in response to normal differentiation cues (Mikkelsen et al. 2007; Squazzo et al. 2006). The preponderance of promoter DNA methylation in such genes for which low but inducible expression is important for the balance between stem cell maintenance and differentiation is an important class of TSGs which does not fit the classic TSG definition but fits the above mentioned refined "continuum" model of TSGs. When silenced by DNA methylation, these genes may not be appropriately reactivated and thus hamper proper differentiation (Fig. 1b). Their silencing by promoter methylation-mediated non-inducible, low expression state may help in the stem-like state of cancers. Key examples of such genes are the developmental and differentiation regulators, such as *SFRPs* and *SOX17*, that are almost never mutated but are hypermethylated in many cancers and may act as TSGs by antagonizing Wnt-signaling. We recently showed that silencing these genes by themselves are not

the driver events, but their inactivation causes the necessary and sufficient defects in differentiation that creates a permissive state for tumorigenesis by cancer driver mutations, such as oncogenic BRAF (Tao et al. 2019). Increased stem cell maintenance and decreased or abnormal differentiation potential of cancers is thus an important outcome of such abnormal epigenetic silencing. Detailed characterization of the function of CGI methylated genes during tumorigenesis will provide insights into early tumorigenesis and potential of using these as early biomarkers of cancer promoting defects in normal differentiation. A key challenge in understanding the various roles of these genes in tumor development will be to use an array of experimental approaches, involving in vitro, ex vivo and mouse models to characterize the roles of these genes. Further, DNA methylation alterations track with the type of driver mutation in a cancer, indicating that roles of these genes should be analyzed in the context of the key driver mutations and the tumor signaling pathways involved.

3.3 Structurally Ordered Genomic Domains and Distal Regulatory Elements

The human genome is organized into modular domains called "topologically associating domains" (TADs) delineated by discrete boundaries of insulator elements. TADs, which range in the megabase-scale, are observed with specialized techniques like HiC that map proximity of stretches of distant genomic elements, thus providing a spatial organization of chromatin in the nucleus (Dixon et al. 2016). Through structured organization of chromatin, distal regulatory elements (enhancers) in the genome interact in a specific and regulated manner to orchestrate regulation of gene expression. Specialized proteins bind to the boundaries (insulator elements) of the TADs and help maintain the discreteness of these domains by acting as borders to limit spreading of chromatin factors and histone modifications. The organization of chromosomes into TADs facilitates physical proximity of chromatin elements that otherwise are localized over considerable distances on the linear DNA. Chromosome looping within the TADs brings promoters and distal regulatory elements in close proximity and their interaction is mediated by transcription factors such as the mediator complex, cohesins and CTCF. Looping itself is governed by insulator binding proteins such as CTCF (Rao et al. 2014). Such 3D-dimensional organization of the chromatin plays important roles in regulating gene expression during differentiation and development as indicated by the fact that tissue-specific genes are frequently in association with distal enhancers in a cell-type specific manner. CRISPR-mediated re-engineering in mice of a chromosomal rearrangement, which in human patient families disrupts a TAD locus associated with limb malformations, severely phenocopies the human condition in mice (Lupianez et al. 2015). Thus the mere disruption of TADs results in profound alteration of gene expression. The underlying DNA sequence in insulator elements at the boundaries of TADs are

subject to DNA methylation, which has profound implications on their activity via altering the binding of proteins, such as CTCF, which are essential for maintaining TADs. In cancers, given the genome wide alterations of DNA methylation that can be linked to genome instability and rearrangements, there are profound implications resulting from disruption of TADs, and thereby the regulatory elements, as is being discovered and appreciated in recent years and discussed below.

An important class of the distal regulatory elements is the enhancers making up ~10% of the human genome, which are cell type specific genomic sequences that regulate transcriptional activity of genes that are thousands to million bases away (Jin et al. 2013; Li et al. 2012; Rao et al. 2014). Although enhancers are mainly regulated at the level of TF binding, chromatin composition and histone modifications, and some of these features distinguishes them from gene promoters, (Heintzman et al. 2007; Kim and Shiekhattar 2015; Shlyueva et al. 2014), there is increasing evidence that activity of enhancers is linked to, and may also be regulated by, DNA methylation (Jones 2012). Activity of enhancers is tissue type specific, and importantly CpG methylation in enhancers is also observed to be tissue type specific. In genomewide analyses of tissue type specific methylation patterns, about 26% of the cell type specific DNA methylation sites overlap with putative enhancers, and another 40% of such sites overlap with DNAse I hypersensitive sites which are features of regulatory elements (Ziller et al. 2013). About 90% of hypomethylated regions in colon cancer compared to normal colon contain enhancers (Berman et al. 2011). Importantly, DNA methylation changes at enhancers are better correlated to expression changes of target genes than promoter-CGI methylation (Aran et al. 2013; Leadem et al. 2018). This latter observation is due to the fact that CGI methylation is rare in normal non-neural somatic cells, wherein expression of genes is regulated by chromatin modifications at the CGI promoters and rarely by methylation alteration of promoter CGI (Baylin and Jones 2016; Suzuki and Bird 2008). Majority of enhancers have low density of CpG dinucelotides, which when methylated are associated with absence of the active enhancer histone mark (H3K4me1). Thus it is suggested that DNA methylation state of enhancer can also direct histone modifications (Fig. 2), and play deterministic roles in the activity of enhancers. Unmethylated or partially methylated enhancers can recruit a class of histone methyltransferases that catalyze the H3K4me1 mark, thus marking enhancer regions for poised or active state (Sharifi-Zarchi et al. 2017). Alternately, methylation of CTCF sites in insulator elements can alter chromatin looping, allowing interaction of enhancer elements with target gene promoters (Fig. 2). These studies highlight the importance of understanding DNA methylation alterations in enhancers, in addition to the promoter methylation changes, to understand epigenetic deregulation of gene expression in tumor development.

Since enhancer regions are enriched for transcription factor (TF) binding sites, and are regulated by the TFs, DNA methylation alterations will interfere with TF binding and alter regulation of target genes. In concordance with the global hypomethylation in cancers, large-scale analysis of enhancers across multiple tumor types showed that majority of enhancers undergo hypomethylation in cancers. Analyses of array based DNA methylation data, which limits analysis to only a

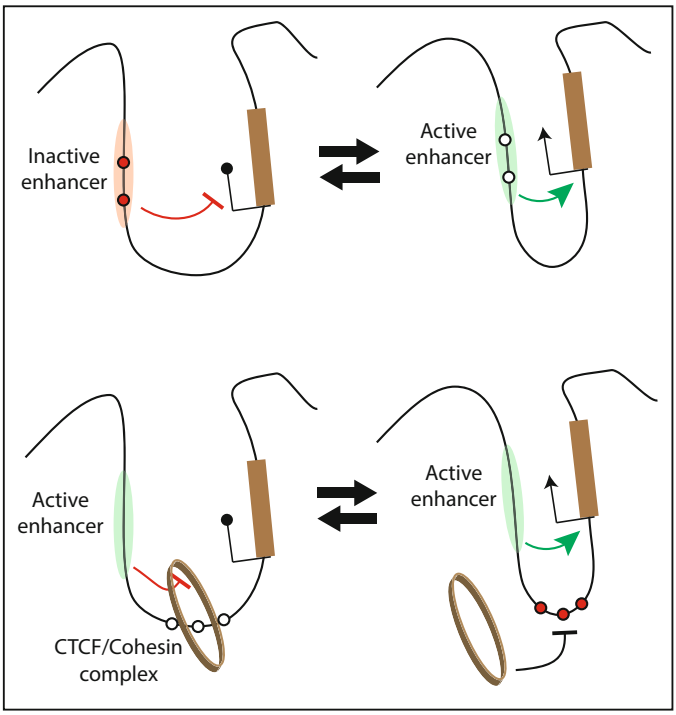

Fig. 2 Distal regulatory elements are subject to regulation by DNA methylation. Methylation alterations in non-promoter regions, aberrant hypermethylation or hypomethylation of enhancers, can cause deregulated activation or inactivation of genes. *Top panel:* Methylation of CpG dinucleotides in enhancer elements causes inactivation of enhancers and prevents activation of target genes, which may be in poised expression state. Hypomethylation of enhancer causes recruitment of DNA binding factors and activating histone modifications, which allows target gene activation. *Bottom panel:* Methylation of CpG sites in insulator regions bound by CTCF/cohesin complexes prevents trans-activation of target gene by enhancers. Altered methylation of the CpG sites at the insulator elements causes interaction of distal enhancers and genes, causing activation of genes

subset of all putative enhancers for which probes are available, showed ~6000 and 1200 enhancers undergo hypo- or hypermethylation, respectively, impacting thousands of genes across various tumors (Yao et al. 2015). As would be expected, hypomethylation of an enhancer is correlated with upregulation of the potential target genes while hypermethylation is associated with downregulation. Interestingly, known tumor suppressor genes (like *CDKN1A*, *SPRY2*) were downregulated in association with the corresponding enhancer methylation while various oncogenes (like *MYC*, *TERT)* were upregulated in association with hypomethylated enhancers. In the context of the PMDs described earlier, enhancer deregulation due to DNA hypomethylation might be pervasive during successive rounds of cell divisions in cancer cells, as well as normal aging cells.

As with enhancers, CTCF sites that mark insulator elements controlling chromosomal looping are under tight regulation by DNA methylation. CTCF binding sites

contain CpG-sites that when methylated abrogate binding of CTCF. Thus, considering the genomewide changes in DNA methylation in cancers, it then only remains a matter of exploration of these regions to understand the extent of their disruption, the effect on the TADs and the ensuing gene expression changes. A key study showed how in IDH mutant cases of gliomas, which are associated with increased CpG-methylation, a CTCF-binding site gets aberrantly methylated disrupting binding of CTCF resulting in abrogation of the tight insulator function. As a result, the oncogene *PDGFRA* is able to interact with an enhancer located ~900 kb away in a neighboring TAD resulting in increased *PDGFRA* expression. Importantly, this enhancer otherwise does not interact with *PDGFRA* in normal and gliomas without the IDH mutation (Flavahan et al. 2016). Such abnormal activation of oncogenes, and silencing of tumor suppressor genes, will probably prove to be a theme across various cancer types because of the large-scale DNA methylation deregulation.

4 Origin of Cancer Methylation Changes from Normal Physiological Processes

The global methylation changes described above in cancers are observed in multiple normal physiological processes involving continued mitotic cycling of somatic cells, such as during tissue regeneration associated with aging, and other processes like inflammation, immortalization and senescence. Continuous cycling of cells during aging involves both global hypomethylation and promoter hypermethylation (Issa 2014). One of the earliest observations in the field is that de novo promoter methylation of ER and IGF2 genes occurs in the colonic mucosa of aged individuals, and methylation of these progressively increased in cancers (Issa 2014). Thereon, multiple gene promoters methylated in cancers have been shown to be also methylated in aging. Global analyses in multiple tissues in mice and human have demonstrated profound genome-wide changes in the DNA methylation levels during aging (Christensen et al. 2009; Hannum et al. 2013; Heyn et al. 2012; Maegawa et al. 2010, 2014, 2017; Rakyan et al. 2010; Teschendorff et al. 2010). Since promoter methylation events in cancers are very early and frequent events, similar methylation patterns observed during aging have suggested a potential for transition of aged cells with altered methylation patterns to tumor initiating cells (Fig. 3). Multiple aspects of age-related methylation, beyond the genes that are specifically affected, compels understanding age-related methylation important for understanding the etiology of cancer. Comparison of methylation alterations occurring during in vitro immortalization and transformation of cells by serial expression of hTERT, SV40 large T and HRAS in relation to cells undergoing senescence showed that the transformation-associated methylation arise stochastically, while senescence-associated methylation arise in a defined and programmatic manner. Importantly, genes that get stochastically methylated during transformation, compared to those specifically methylated during senescence, are more likely to be also methylated in

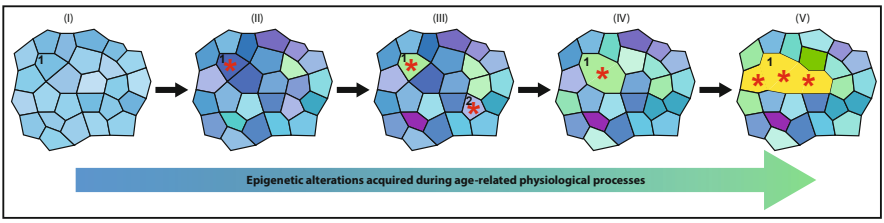

Fig. 3 Progressive epigenetic changes during aging provide permissive state for oncogenic driver effect of pre-existing or acquired mutations. The figure shows schematic of regions in tissues/cell subpopulations harbor related DNA methylation patterns (colored polygons), which diverge with aging. Most likely this methylation heterogeneity is maintained in different long-living stem cells that will give rise to subclonal populations carrying similar epigenetic marks. Oncogenic mutations (region 1) is initially not tumorigenic (I, II) unless a permissive epigenetic state is achieved (III). Mutations occurring in cells with non-permissive epigenetic background (region 2, III) are not transformative and lost. Epigenetic states keep diverging further, and continue to evolve in the tumor cells (IV, V)

aging tissues, suggesting that the methylation changes observed in cancers most likely derive from normally dividing, aging cells (Xie et al. 2018).

Since stem-cells are the long-lived components of aging tissues, alterations in maintenance of DNA methylation are expected to be fixed, and further evolve in the stem cells during continuous prolonged divisions during tissue replenishment and aging (Fig. 3). Compelling evidence for the origin of age-related methylation alterations arising in long-lived continuously dividing stem cells is the contrasting age-related methylation changes in the continuously dividing colon stem cell population vs. the rarely dividing stem cells in the hair follicles (Shibata 2009). Whereas age-related progressive methylation changes are observed in the colonic stem cells, hair follicle stem cells show no such changes during the lifetime of an individual. The observation that many of the genes methylated during aging belong to the PcG targets also methylated in cancer (Rakyan et al. 2010; Teschendorff et al. 2010) may innocuously suggest that such methylation changes in both cancer and aging is just a mere consequence of an inherent bias for these genes to be methylated during multiple mitotic divisions. Although this could be true for many genes by nature of them being PcG marked, the same sets of genes epigenetically altered in aging and cancers, potentially impacting tumor suppressors and stem cell pathways (Easwaran et al. 2012; West et al. 2013), additionally suggest that age-related methylation of these genes may increase cell fitness allowing for clonal expansion and neoplasms to develop (Fig. 3). In regard to this, methylation of CpG residues is highly polymorphic in multiple primary tissues, which increase in tumors and in primary cells in culture (Landan et al. 2012). The diversity in methylation patterns arising from polymorphic methylation is suggested to provide the necessary variation for Darwinian selection of fitter clones (Hansen et al. 2011) (Fig. 3). For example, DNA methylation occurs at key colorectal cancer (CRC) and adenoma related control genes, such as *APC*, *AXIN2*, *DKK1*, *HPP1*, *N33*, *CDKN2A/p16*, *SFRP1*, *SFRP2* and *SFRP4*, during ageing (Belshaw et al. 2008). Some of these genes are otherwise mutated, a key

example being *CDKN2A/p16*, which is a critical gene that prevents oncogene-induced transformation. Preexisting epigenetic silencing of multiple such genes in the same cell, which may occur due to stochastically arising polymorphic methylation patterns during aging, will sufficiently block cell cycle checkpoint and activate stem-cell pathways allowing oncogenic-driver effects. Recently we showed that simultaneous inactivation of genes subject to epigenetic silencing in colon cancers, namely *CDX2*, *SFRP4*, *SOX17* and *CDKN2A*, sufficiently creates the permissive state for oncogenic-BRAF induced transformation of colon derived organoid cultures (Tao et al. 2019).

How cancer-related methylation changes come about to be is tightly linked to understanding tumor initiation. Various age-related physiological processes, like chronic inflammation and carcinogen induced genomic stress, acutely trigger epigenetic changes observed in cancers (Asada et al. 2015; Blanco et al. 2007; Hahn et al. 2008; Niwa and Ushijima 2010; O'Hagan et al. 2011; Vaz et al. 2017). As discussed above, age-related methylation of CGI promoters and other genomic elements (both hypo- and hypermethylation) will help in the initial stages of tumorigenesis, and these methylation patterns may get selected and further expanded during continuous cell divisions (Fig. 3). In the context of the current framework of oncogenesis, sequential occurrence of mutations allow expansion of fitter cells by both neutral evolution and selection causing genetically heterogeneous tumors (McGranahan and Swanton 2017). Undoubtedly random somatic mutations that accumulate during aging are central drivers in this framework of tumorigenesis. However, substantial numbers of somatic mutations in mice occur during the growth phases early in life, and the rate of mutation accumulation slows down once stem cell divisions decrease and as organs enter maintenance mode (Vijg et al. 2005) [discussed in Rozhok and DeGregori (2016)]. Importantly cancer mutations precede tumor incidence by years to decades (Brucher and Jamall 2016; Desai et al. 2018; Forsberg et al. 2013; Mori et al. 2002; Vogelstein et al. 2013), whereas the latter increases exponentially with age. This is contrasted for methylation alterations that continue to deviate and accumulate during aging (Fraga et al. 2005; Hannum et al. 2013). Thus in the current framework of oncogenesis, a key question regarding the age-related cancer risk is what non-genetic factors allow the impact of driver mutations to unfold as a function of age. Progressive age associated epigenetic modifications may be one of the key factors in this (Fig 3), and for most age-related cancer incidences, methylation patterns observed in cancers may thus originate from subpopulations of aging cells that carry epigenetic alterations that creates a permissive state for tumor initiation (Fig. 3).

5 Molecular Mechanisms of Methylation Patterning

Establishment and maintenance of DNA methylation relies on combined action of the three major DNA methyltransferase, DNMT1, DNMT3A and B. The latter two enzymes have been proposed to mainly play roles in de novo DNA methylation. During embryonic development in mice, Dnmt3a is implicated in establishing

methylation at imprinted genes discriminating genes by parent of origin (Kaneda et al. 2004; Okano et al. 1999) while Dnmt3b is involved in methylation of pericentromric repeats (Okano et al. 1999; Xu et al. 1999). The de novo methylation activity of Dnmt3 enzymes is directed to defined chromatin regions by specialized protein sequences in its N-terminal domain. DNMT3A, and possibly DNMT3B too, exist in an auto-inhibitory inactive state that specifically is activated by direct interaction with H3 histone tail unmethylated at lysine-4 (H3K4me0) (Guo et al. 2015). Various active chromatin modifications, including H3K4me3, inhibit interaction of DNMT3 with the H3-tail. In CGI promoters, the H3K4me3 mark is enriched by targeting the H3K4 methyltransferase SET1 complex (MLL1/2) via specific binding of the CXXC-domain containing CFP1 to unmethylated CpGs (Baubec and Schubeler 2014; Clouaire et al. 2014; Thomson et al. 2010). Hence the underlying CpG density of CGI promoters itself attracts H3K4me3 marks, which in turn antagonizes methylation by the DNMT3 enzymes, thus maintaining CGI promoters in an unmethylated state. In contrast to the DNMT3-chromatin interactions ensuring methylation free zone at CGI, the PWWP domain of DNMT3 directs the enzyme to H3K36me2/3 regions of the chromatin, which corresponds to gene bodies and heterochromatin (Dhayalan et al. 2010; Zhang et al. 2010). Further, DNMT3A/B interaction with HP1 and G9A recruits it to H3K9me3 residues in pericentric heterochromatin (Lehnertz et al. 2003) and gene promoters (Epsztejn-Litman et al. 2008) respectively. Additional control of the DNMT3 activity towards unmethylated DNA is imposed by secluding DNMT3 to the methylated DNA/chromatin in heterochromatic fraction, and degradation of free-floating enzymes (Jeong et al. 2009; Sharma et al. 2011). In somatic cells, where the expression of DNMT3 is markedly reduced, via the above mechanisms of recruitment to non-H3K4me3-containing, CpG methylated chromatin, the activity of DNMT3 enzymes is restrained to already methylated regions where they cooperate with DNMT1 in maintaining the methylation patterns (Jeltsch and Jurkowska 2014; Jones and Liang 2009; Sharma et al. 2011). The inherent ability of DNMT3 enzymes to read the 'histone code' thus seems to be a prominent mechanism in establishing methylation patterns during development. The established methylation patterns are then maintained during successive rounds of replication mainly by DNMT1, but also by the DNMT3 enzymes by virtue of their affinity for methylated CpG containing nucleosomes (Jones and Liang 2009).

In line with the above paradigm for establishment of CpG methylation patterns by the histone code, the underlying DNA sequence has an important role in determining if regions of the genome will or will not be methylated (Lienert et al. 2011; Stadler et al. 2011). This role of the DNA sequence may precede or work in parallel with the histone code. For example, cis-acting sequences (~700–1000 bp) that have affinity to DNA binding transcription factors protect DNA from methylation independent of transcriptional activity or CpG density (Lienert et al. 2011). Importantly, the cis-acting sequences containing TF binding sites can protect exogenous DNA from CpG methylation, as well as can cause demethylation of exogenously introduced methylated DNA. In this model, TF binding per se may sterically hinder access of DNA methyltransferases to DNA, or more likely the TF could mediate

recruitment of the MLL class of histone methyltransferases leading to H3K4me3, which as described above will prevent de novo CpG methylation (Demers et al. 2007; Rao and Dou 2015). Further, as described earlier, silencing of CGI promoters during normal development and differentiation is mediated by H3K27me3 mark, which is regulated by the polycomb repressive complex (PRC) 1 and 2. Presence of the H3K27me3 mark is normally anti-correlated with DNA methylation in normal and cancer cells (Easwaran et al. 2012; Kouzarides 2007; Sen et al. 2016). Accordingly the polycomb mark has been linked to preventing DNA from methylation. A component of the PRC1 complex, FBXL10/KDM2B, containing the CXXC domain that can bind to unmethylated CpG-rich sequence, plays the primary role in this anti-correlation between DNA methylation and H3K27me3. Specifically, in ES cells FBXL10/KDM2B occupies ~90% of all promoter-associated CGI, and an equal proportion of CGI promoters that are silenced by PRC1/2 complexes (Farcas et al. 2012; He et al. 2013; Wu et al. 2013). FBXL10 prevents CpG methylation by binding to the PRC regulated CGI promoters, as loss of FBXL10 results in rapid methylation of only those promoters that are polycomb regulated (Boulard et al. 2015). It is important to note that, as introduced earlier, polycomb-regulated genes are more often methylated in cancers. Deregulation of FBXL10 mechanics in cancers, for example it is mutated in diffuse B-cell lymphomas (Pasqualucci et al. 2011), may cause aberrant promoter CGI methylation for some of the genes during aging and tumorigenesis.

In summary, the cellular transcription program established during development and differentiation, as defined by the expressed repertoire of transcription and DNA binding factors, sufficiently is able to establish and maintain the DNA methylation epigenetic program (Burger et al. 2013; Stadler et al. 2011). In this model, absence of DNA binding factors to cis-elements may trigger DNA methylation, and presence will protect from DNA methylation. The specific molecular details of the dynamics of TF binding, histone modifications and recruitment of DNMTs in mediating methylation patterns still needs elucidation. A noteworthy aspect from the above discussion is that modes of both gene activation and silencing of promoter CGIs, viz. by H3K4me3 and H3K27me3 respectively have central roles in keeping DNA methylation at bay. And the evidence points that there is a concerted role of underlying DNA sequence (TF binding sites and unmethylated CpG dense regions) directed histone code in establishing the methylation patterns. Although mechanics of the CpG methylation patterning is better worked out for the CGI promoters, the remainder of genome may permit CpG methylation due to lack of such DNA sequences and the specific activating histone marks.

6 Road to Cancer Methylome

A key question is how the normal mechanics of DNA methylation establishment and maintenance are perturbed in cancers, and its precursor aging cells. Above discussions indicate that there are layers of molecular deregulation in response to external

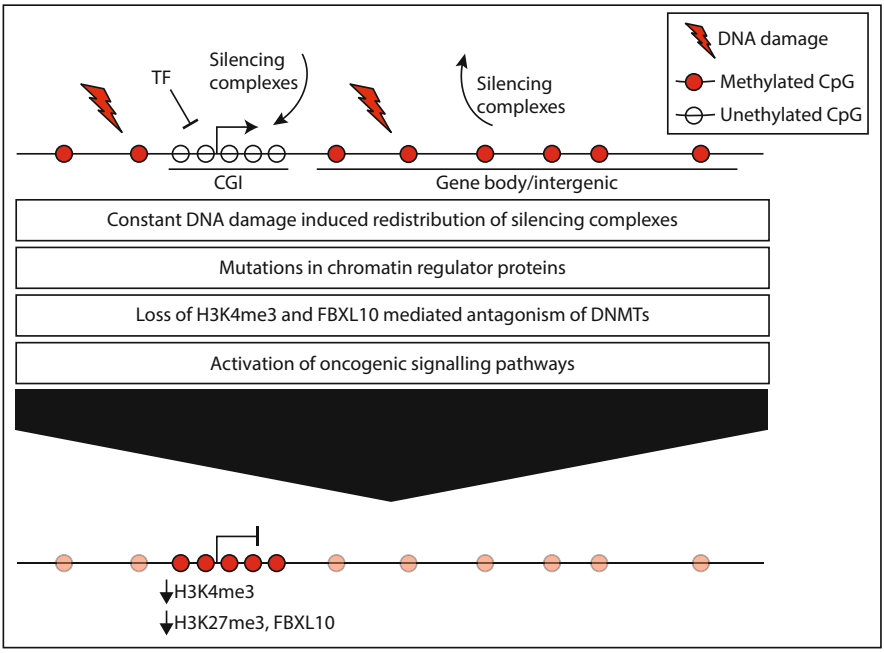

Fig. 4 Molecular basis for altered methylation in cancers. Multiple layers of deregulation of chromatin proteins mediate methylation alterations during aging and tumorigenesis. These alterations in general result in changes in recruitment of silencing complexes to CGI promoters, and their simultaneous loss from gene body and intergenic regions

stimuli (exposure to carcinogenic genotoxic stressors), microenvironmental chemokine changes (inflammation, aging), cell intrinsic pathway alterations (oncogenic and epigenetic modifier mutations), that in conjunction with selection of methylation profiles lead to the ultimate methylation landscape in cancers. Individual layers in this regulation are as follows (Fig. 4):

(a) DNA damage causes genomewide chromatin changes involving altered recruitment of DNMTs, SIRT1, PRC components, CHD4 (NuRD silencing complex component) to damaged CGI promoters, which helps maintain repressed chromatin at CGI promoters. At the same time, DNMTs and SIRT1 is released from the remainder of the genome, potentially causing hypomethylation. Thus continuous DNA damage caused due to cell intrinsic ROS levels, environmental chemical exposure, inflammatory microenvironment and oncogenic stress will lead to gradual genomewide aberrations in DNA methylation (O'Hagan et al. 2011; Vaz et al. 2017). Particularly, these processes can be key drivers of stochastic methylation changes during aging and inflammation.
(b) Direct loss- and gain-of-function mutations in chromatin regulator proteins will impact the methylome. In this regard, mutations of chromatin proteins are the most common class of cancer mutations, and are observed across various cancers

(Dawson 2017; Shen and Laird 2013). For majority of these mutations, their roles in modulating the methylome need to be investigated. Leading examples of mutations that directly impact methylation are the IDH and TET enzymes. IDH1 and 2 are metabolic enzymes that are not directly involved in chromatin modulation, but their mutation in cancers causes neomorphic enzyme activity that converts α-ketoglutarate (α-KG) to D-2-hydroxyglutarate (D-2-HG). The latter is an oncometabolite that inhibits various Fe(II)/2-oxoglutarate-dependent dioxygenases (Dang et al. 2009; Xu et al. 2011). These include various histone demethylases that protect against DNA methylation by diminishing chromatin marks that attract DNA methylation, and the TET family of enzymes that catalyze DNA demethylation by oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC). This has directly been implicated in increased DNA methylation phenotype, called the CpG-island methylator phenotype, in acute myeloid leukemia (AML) (Figueroa et al. 2010) and gliomas (Noushmehr et al. 2010; Turcan et al. 2012). On the other hand, enzymes involved in DNA demethylation discussed earlier, like *TET2*, are also mutated in various cancers, especially prominent in the hematological malignancies (Scourzic et al. 2015). *TET2* mutations majorly cause methylation of distal regulatory enhancer elements (Rasmussen et al. 2015).

(c) Cancers involve significant expression changes in transcription factor repertoire, a striking example being the almost universally overexpressed MYC factor in various cancers (Dang 2012). This will alter occupancy of TFs at CGI promoters and distal regulatory elements. Altered presence or absence of TFs will cause changes to the histone code thereby causing changes to DNA methylation patterns (Gebhard et al. 2010). Redistribution of DNA binding factors can be directly linked to oncogenic mutations, such as the frequent MEK-ERK pathway activating mutations in RAS/RAF. In ES cells, inhibition of MEK, has been shown to reduce global DNA methylation by reducing DNMT3 enzymes and activating TET1 (Sim et al. 2017). Similarly, growth-factor signalling pathways, such as FGF and Wnt that are activated in many epithelial cancers, have been proposed to induce PRC2-dependent CGI methylation. These latter implications have been made from investigations on the methylation dynamics occurring in extraembryonic tissue of the trophoectoderm lineage, which has revealed de novo promoter CGI methylation of the same developmental regulator genes methylated in cancers (Smith et al. 2017). In these studies, signalling by FGF and WNT increases the promoter CGI methylation. This conservation of CGI methylation patterns in the extraembryonic ectoderm and human cancers indicate strong parallels in activation of oncogenic signalling pathways and PRC-dependent CGI methylation. In concordance with these latter studies, we showed recently that continued culturing, over several months, of colon derived organoids in Wnt-enriched medium used for organoid growth results in an aging-like accumulation of DNA methylation at key developmental regulator promoters, which is necessary for oncogenic-BRAF induced transformation (Tao et al. 2019).

(d) The observations that PcG-regulated promoters are most prone to get methylated indicate a role for the PcG components in mediating recruitment of DNMTs. A key finding supporting this is that loss of FBXL10/KDM2B causes methylation of only those promoters associated with PRC1/2 (Boulard et al. 2015). Whether, and how, loss of FBXL10 from some PcG occupied promoters occurs during tumor development remains to be studied. In this regard, how PRC-mediated recruitment of DNMTs occurs in the context of H3K4me3 and H3K27me3 marked bivalent chromatin promoters is important to understand. Presence of active or stalled RNA Pol II, which occur with H3K4me3, is associated with protection from CGI methylation while presence of H3K27me3 predisposes to CGI methylation (Takeshima et al. 2009). Thus, a sequential step may involve removal of transcriptionally poised state to a PcG regulated promoter, which subsequently may acquire methylation by altered activity of FBXL10.

7 Conclusions and Future Directions

Accumulation of DNA methylation alterations occurs during various normal processes, importantly during aging. Increasing evidence suggests that these alterations have a role in predisposing to tumorigenesis. How all the mechanisms suggested above interlace to produce the epigenetic drift during aging, and in cancer, is important to understand the specific means by which various cancer predisposition factors work through modulating the epigenome. Equally important is to understand the role of the epigenetically modified genes during the early steps of tumorigenesis. We predict that development of appropriate markers that can differentiate epigenetically altered cell populations, in response to intrinsic (such as inflammation) and extrinsic (such as environmental exposures) factors, carrying aberrant methylation of functionally relevant genes holds promise in identifying cancer risk states.

References

Almouzni G, Cedar H (2016) Maintenance of epigenetic information. Cold Spring Harb Perspect Biol 8. pii: a019372

Aran D, Sabato S, Hellman A (2013) DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. Genome Biol 14:R21

Asada K, Nakajima T, Shimazu T et al (2015) Demonstration of the usefulness of epigenetic cancer risk prediction by a multicentre prospective cohort study. Gut 64:388–396

Baubec T, Schubeler D (2014) Genomic patterns and context specific interpretation of DNA methylation. Curr Opin Genet Dev 25:85–92

Baubec T, Ivanek R, Lienert F et al (2013) Methylation-dependent and -independent genomic targeting principles of the MBD protein family. Cell 153:480–492

Baylin SB, Jones PA (2016) Epigenetic determinants of cancer. Cold Spring Harb Perspect Biol 8. pii: a019505

Belshaw NJ, Elliott GO, Foxall RJ et al (2008) Profiling CpG island field methylation in both morphologically normal and neoplastic human colonic mucosa. Br J Cancer 99:136–142
Berger AH, Knudson AG, Pandolfi PP (2011) A continuum model for tumour suppression. Nature 476:163–169
Bergman Y, Cedar H (2013) DNA methylation dynamics in health and disease. Nat Struct Mol Biol 20:274–281
Berman BP, Weisenberger DJ, Aman JF et al (2011) Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. Nat Genet 44:40–46
Bernstein BE, Mikkelsen TS, Xie X et al (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125:315–326
Bird A (2007) Perceptions of epigenetics. Nature 447:396–398
Blanco D, Vicent S, Fraga MF et al (2007) Molecular analysis of a multistep lung cancer model induced by chronic inflammation reveals epigenetic regulation of p16 and activation of the DNA damage response pathway. Neoplasia 9:840–852
Boulard M, Edwards JR, Bestor TH (2015) FBXL10 protects Polycomb-bound genes from hypermethylation. Nat Genet 47:479–485
Brucher BL, Jamall IS (2016) Somatic mutation theory – why it's wrong for most cancers. Cell Physiol Biochem 38:1663–1680
Burger L, Gaidatzis D, Schubeler D et al (2013) Identification of active regulatory regions from DNA methylation data. Nucleic Acids Res 41:e155
Chen RZ, Pettersson U, Beard C et al (1998) DNA hypomethylation leads to elevated mutation rates. Nature 395:89–93
Christensen BC, Houseman EA, Marsit CJ et al (2009) Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet 5:e1000602
Chuang LS, Ian HI, Koh TW et al (1997) Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science 277:1996–2000
Clouaire T, Webb S, Bird A (2014) Cfp1 is required for gene expression-dependent H3K4 trimethylation and H3K9 acetylation in embryonic stem cells. Genome Biol 15:451
Dang CV (2012) MYC on the path to cancer. Cell 149:22–35
Dang L, White DW, Gross S et al (2009) Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature 462:739–744
Dawson MA (2017) The cancer epigenome: concepts, challenges, and therapeutic opportunities. Science 355:1147–1152
Deaton AM, Bird A (2011) CpG islands and the regulation of transcription. Genes Dev 25:1010–1022
Demers C, Chaturvedi CP, Ranish JA et al (2007) Activator-mediated recruitment of the MLL2 methyltransferase complex to the beta-globin locus. Mol Cell 27:573–584
Desai P, Mencia-Trinchant N, Savenkov O et al (2018) Somatic mutations precede acute myeloid leukemia years before diagnosis. Nat Med 24:1015–1023
Dhayalan A, Rajavelu A, Rathert P et al (2010) The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. J Biol Chem 285:26114–26120
Dixon JR, Gorkin DU, Ren B (2016) Chromatin domains: the unit of chromosome organization. Mol Cell 62:668–680
Easwaran H, Schermelleh L, Leonhardt H et al (2004) Replication-independent chromatin loading of Dnmt1 during G2 and M phases. EMBO Rep 5:1181–1186
Easwaran H, Leonhardt H, Cardoso M (2005) Cell cycle markers for live cell analyses. Cell Cycle 4:453–455
Easwaran H, Johnstone SE, Van Neste L et al (2012) A DNA hypermethylation module for the stem/progenitor cell signature of cancer. Genome Res 22:837–849
Eden A, Gaudet F, Waghmare A et al (2003) Chromosomal instability and tumors promoted by DNA hypomethylation. Science 300:455

Ehrlich M, Lacey M (2013) DNA hypomethylation and hemimethylation in cancer. Adv Exp Med Biol 754:31–56
Epsztejn-Litman S, Feldman N, Abu-Remaileh M et al (2008) De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. Nat Struct Mol Biol 15:1176–1183
Farcas AM, Blackledge NP, Sudbery I et al (2012) KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. elife 1:e00205
Figueroa ME, Abdel-Wahab O, Lu C et al (2010) Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell 18:553–567
Flavahan WA, Drier Y, Liau BB et al (2016) Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature 529:110–114
Forsberg LA, Absher D, Dumanski JP (2013) Republished: Non-heritable genetics of human disease: spotlight on post-zygotic genetic variation acquired during lifetime. Postgrad Med J 89:417–426
Fraga MF, Ballestar E, Paz MF et al (2005) Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A 102:10604–10609
Gao Y, Teschendorff AE (2017) Epigenetic and genetic deregulation in cancer target distinct signaling pathway domains. Nucleic Acids Res 45:583–596
Gaudet F, Hodgson JG, Eden A et al (2003) Induction of tumors in mice by genomic hypomethylation. Science 300:489–492
Gebhard C, Benner C, Ehrich M et al (2010) General transcription factor binding at CpG islands in normal cells correlates with resistance to de novo DNA methylation in cancer cells. Cancer Res 70:1398–1407
Guo X, Wang L, Li J et al (2015) Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. Nature 517:640–644
Hahn MA, Hahn T, Lee DH et al (2008) Methylation of polycomb target genes in intestinal cancer is mediated by inflammation. Cancer Res 68:10280–10289
Han H, Cortez CC, Yang X et al (2011) DNA methylation directly silences genes with non-CpG island promoters and establishes a nucleosome occupied promoter. Hum Mol Genet 20:4299–4310
Hannum G, Guinney J, Zhao L et al (2013) Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell 49:359–367
Hansen KD, Timp W, Bravo HC et al (2011) Increased methylation variation in epigenetic domains across cancer types. Nat Genet 43:768–775
Hartung T, Zhang L, Kanwar R et al (2012) Diametrically opposite methylome-transcriptome relationships in high- and low-CpG promoter genes in postmitotic neural rat tissue. Epigenetics 7:421–428
He J, Shen L, Wan M et al (2013) Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. Nat Cell Biol 15:373–384
Heintzman ND, Stuart RK, Hon G et al (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet 39:311–318
Herman JG, Baylin SB (2003) Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med 349:2042–2054
Heyn H, Li N, Ferreira HJ et al (2012) Distinct DNA methylomes of newborns and centenarians. Proc Natl Acad Sci U S A 109:10522–10527
Holm TM, Jackson-Grusby L, Brambrink T et al (2005) Global loss of imprinting leads to widespread tumorigenesis in adult mice. Cancer Cell 8:275–285
Hur K, Cejas P, Feliu J et al (2014) Hypomethylation of long interspersed nuclear element-1 (LINE-1) leads to activation of proto-oncogenes in human colorectal cancer metastasis. Gut 63:635–646
Illingworth RS, Bird AP (2009) CpG islands – a rough guide. FEBS Lett 583:1713–1720
Issa JP (2014) Aging and epigenetic drift: a vicious cycle. J Clin Invest 124:24–29

Jackson DA, Pombo A (1998) Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. J Cell Biol 140:1285–1295
Jeltsch A, Jurkowska RZ (2014) New concepts in DNA methylation. Trends Biochem Sci 39:310–318
Jeong S, Liang G, Sharma S et al (2009) Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. Mol Cell Biol 29:5366–5376
Jin F, Li Y, Dixon JR et al (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature 503:290–294
Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet 13:484–492
Jones PA, Liang G (2009) Rethinking how DNA methylation patterns are maintained. Nat Rev Genet 10:805–811
Kaneda M, Okano M, Hata K et al (2004) Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature 429:900–903
Kim TK, Shiekhattar R (2015) Architectural and functional commonalities between enhancers and promoters. Cell 162:948–959
Knudson AG (2001) Two genetic hits (more or less) to cancer. Nat Rev Cancer 1:157–162
Kouzarides T (2007) Chromatin modifications and their function. Cell 128:693–705
Landan G, Cohen NM, Mukamel Z et al (2012) Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues. Nat Genet 44:1207–1214
Lawrence MS, Stojanov P, Polak P et al (2013) Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 499:214–218
Leadem BR, Kagiampakis I, Wilson C et al (2018) A KDM5 inhibitor increases global H3K4 trimethylation occupancy and enhances the biological efficacy of 5-Aza-2'-deoxycytidine. Cancer Res 78:1127–1139
Lehnertz B, Ueda Y, Derijck AA et al (2003) Suv39h-mediated histone h3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol 13:1192–1200
Leonhardt H, Page AW, Weier HU et al (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell 71:865–873
Li G, Ruan X, Auerbach RK et al (2012) Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell 148:84–98
Lienert F, Wirbelauer C, Som I et al (2011) Identification of genetic elements that autonomously determine DNA methylation states. Nat Genet 43:1091–1097
Lister R, Pelizzola M, Kida YS et al (2011) Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature 471:68–73
Lupianez DG, Kraft K, Heinrich V et al (2015) Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell 161:1012–1025
Mack SC, Witt H, Piro RM et al (2014) Epigenomic alterations define lethal CIMP-positive ependymomas of infancy. Nature 506:445–450
Maegawa S, Hinkal G, Kim HS et al (2010) Widespread and tissue specific age-related DNA methylation changes in mice. Genome Res 20:332–340
Maegawa S, Gough SM, Watanabe-Okochi N et al (2014) Age-related epigenetic drift in the pathogenesis of MDS and AML. Genome Res 24:580–591
Maegawa S, Lu Y, Tahara T et al (2017) Caloric restriction delays age-related methylation drift. Nat Commun 8:539
Mazor T, Pankov A, Johnson BE et al (2015) DNA methylation and somatic mutations converge on the cell cycle and define similar evolutionary histories in brain tumors. Cancer Cell 28:307–317
McGranahan N, Swanton C (2017) Clonal heterogeneity and tumor evolution: past, present, and the future. Cell 168:613–628

McKenna ES, Roberts CW (2009) Epigenetics and cancer without genomic instability. Cell Cycle 8:23–26
Mikkelsen TS, Ku M, Jaffe DB et al (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448:553–560
Mori H, Colman SM, Xiao Z et al (2002) Chromosome translocations and covert leukemic clones are generated during normal fetal development. Proc Natl Acad Sci U S A 99:8242–8247
Niwa T, Ushijima T (2010) Induction of epigenetic alterations by chronic inflammation and its significance on carcinogenesis. Adv Genet 71:41–56
Noushmehr H, Weisenberger DJ, Diefes K et al (2010) Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. Cancer Cell 17:510–522
O'Donnell AH, Edwards JR, Rollins RA et al (2014) Methylation abnormalities in mammary carcinoma: the methylation suicide hypothesis. J Cancer Ther 5:1311–1324
O'Hagan HM, Wang W, Sen S et al (2011) Oxidative damage targets complexes containing DNA methyltransferases, SIRT1, and polycomb members to promoter CpG Islands. Cancer Cell 20:606–619
Ohm JE, McGarvey KM, Yu X et al (2007) A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet 39:237–242
Okano M, Bell DW, Haber DA et al (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99:247–257
Pasqualucci L, Trifonov V, Fabbri G et al (2011) Analysis of the coding genome of diffuse large B-cell lymphoma. Nat Genet 43:830–837
Pradhan S, Bacolla A, Wells RD et al (1999) Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. J Biol Chem 274:33002–33010
Rakyan VK, Down TA, Maslau S et al (2010) Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. Genome Res 20:434–439
Rao RC, Dou Y (2015) Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. Nat Rev Cancer 15:334–346
Rao SS, Huntley MH, Durand NC et al (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159:1665–1680
Rasmussen KD, Jia G, Johansen JV et al (2015) Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. Genes Dev 29:910–922
Rozhok AI, DeGregori J (2016) The evolution of lifespan and age-dependent cancer risk. Trends Cancer 2:552–560
Schermelleh L, Haemmer A, Spada F et al (2007) Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. Nucleic Acids Res 35:4301–4312
Schlesinger Y, Straussman R, Keshet I et al (2007) Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet 39:232–236
Schuebel KE, Chen W, Cope L et al (2007) Comparing the DNA hypermethylome with gene mutations in human colorectal cancer. PLoS Genet 3:1709–1723
Scourzic L, Mouly E, Bernard OA (2015) TET proteins and the control of cytosine demethylation in cancer. Genome Med 7:9
Sen S, Block KF, Pasini A et al (2016) Genome-wide positioning of bivalent mononucleosomes. BMC Med Genet 9:60
Sharif J, Muto M, Takebayashi S et al (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature 450:908–912
Sharifi-Zarchi A, Gerovska D, Adachi K et al (2017) DNA methylation regulates discrimination of enhancers from promoters through a H3K4me1-H3K4me3 seesaw mechanism. BMC Genomics 18:964

Sharma S, De Carvalho DD, Jeong S et al (2011) Nucleosomes containing methylated DNA stabilize DNA methyltransferases 3A/3B and ensure faithful epigenetic inheritance. PLoS Genet 7:e1001286
Shen H, Laird PW (2013) Interplay between the cancer genome and epigenome. Cell 153:38–55
Shibata D (2009) Inferring human stem cell behaviour from epigenetic drift. J Pathol 217:199–205
Shlyueva D, Stampfel G, Stark A (2014) Transcriptional enhancers: from properties to genome-wide predictions. Nat Rev Genet 15:272–286
Sim YJ, Kim MS, Nayfeh A et al (2017) 2i maintains a naive ground state in ESCs through two distinct epigenetic mechanisms. Stem Cell Reports 8:1312–1328
Smith ZD, Shi J, Gu H et al (2017) Epigenetic restriction of extraembryonic lineages mirrors the somatic transition to cancer. Nature 549:543–547
Squazzo SL, O'Geen H, Komashko VM et al (2006) Suz12 binds to silenced regions of the genome in a cell-type-specific manner. Genome Res 16:890–900
Stadler MB, Murr R, Burger L et al (2011) DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature 480:490–495
Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 9:465–476
Takai D, Jones PA (2002) Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci U S A 99:3740–3745
Takeshima H, Yamashita S, Shimazu T et al (2009) The presence of RNA polymerase II, active or stalled, predicts epigenetic fate of promoter CpG islands. Genome Res 19:1974–1982
Tao Y, Kang B, Petkovich DA et al (2019) Aging-like spontaneous epigenetic silencing facilitates Wnt activation, stemness and BrafV600E-induced tumorigenesis. Cancer Cell 35(2):315–328
Teschendorff AE, Menon U, Gentry-Maharaj A et al (2010) Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. Genome Res 20:440–446
Thomson JP, Skene PJ, Selfridge J et al (2010) CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature 464:1082–1086
Turcan S, Rohle D, Goenka A et al (2012) IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature 483:479–483
Van Tongelen A, Loriot A, De Smet C (2017) Oncogenic roles of DNA hypomethylation through the activation of cancer-germline genes. Cancer Lett 396:130–137
Vaz M, Hwang SY, Kagiampakis I et al (2017) Chronic cigarette smoke-induced epigenomic changes precede sensitization of bronchial epithelial cells to single-step transformation by KRAS mutations. Cancer Cell 32:360–376 e366
Vijg J, Busuttil RA, Bahar R et al (2005) Aging and genome maintenance. Ann N Y Acad Sci 1055:35–47
Vogelstein B, Papadopoulos N, Velculescu VE et al (2013) Cancer genome landscapes. Science 339:1546–1558
Weber M, Hellmann I, Stadler MB et al (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 39:457–466
West J, Beck S, Wang X et al (2013) An integrative network algorithm identifies age-associated differential methylation interactome hotspots targeting stem-cell differentiation pathways. Sci Rep 3:1630
Widschwendter M, Fiegl H, Egle D et al (2007) Epigenetic stem cell signature in cancer. Nat Genet 39:157–158
Wrangle J, Wang W, Koch A et al (2013) Alterations of immune response of non-small cell lung cancer with azacytidine. Oncotarget 4:2067–2079
Wu H, Zhang Y (2014) Reversing DNA methylation: mechanisms, genomics, and biological functions. Cell 156:45–68
Wu X, Johansen JV, Helin K (2013) Fbxl10/Kdm2b recruits polycomb repressive complex 1 to CpG islands and regulates H2A ubiquitylation. Mol Cell 49:1134–1146
Xie W, Kagiampakis I, Pan L et al (2018) DNA methylation patterns separate senescence from transformation potential and indicate cancer risk. Cancer Cell 33:309–321 e305

Xu GL, Bestor TH, Bourc'his D et al (1999) Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature 402:187–191

Xu W, Yang H, Liu Y et al (2011) Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. Cancer Cell 19:17–30

Yamada Y, Jackson-Grusby L, Linhart H et al (2005) Opposing effects of DNA hypomethylation on intestinal and liver carcinogenesis. Proc Natl Acad Sci U S A 102:13580–13585

Yao L, Shen H, Laird PW et al (2015) Inferring regulatory element landscapes and transcription factor networks from cancer methylomes. Genome Biol 16:105

Zhang Y, Jurkowska R, Soeroes S et al (2010) Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. Nucleic Acids Res 38:4246–4253

Zhou W, Dinh HQ, Ramjan Z et al (2018) DNA methylation loss in late-replicating domains is linked to mitotic cell division. Nat Genet 50:591–602

Ziller MJ, Gu H, Muller F et al (2013) Charting a dynamic DNA methylation landscape of the human genome. Nature 500:477–481