

## Review Article

# The DNA methylation landscape in cancer

Ksenia Skvortsova<sup>1,2</sup>, Clare Stirzaker<sup>1,2</sup> and  Philippa Taberlay<sup>3</sup>

<sup>1</sup>Genomics and Epigenetics Division, Garvan Institute of Medical Research, Sydney 2010, New South Wales, Australia; <sup>2</sup>St Vincent's Clinical School, UNSW Australia, Sydney 2010, New South Wales, Australia; <sup>3</sup>School of Medicine, College of Health and Medicine, University of Tasmania, Hobart 7001, Tasmania, Australia

**Correspondence:** Clare Stirzaker (c.stirzaker@garvan.org.au) or Philippa Taberlay (Phillippa.Taberlay@utas.edu.au)



As one of the most abundant and well-studied epigenetic modifications, DNA methylation plays an essential role in normal development and cellular biology. Global alterations to the DNA methylation landscape contribute to alterations in the transcriptome and deregulation of cellular pathways. Indeed, improved methods to study DNA methylation patterning and dynamics at base pair resolution and across individual DNA molecules on a genome-wide scale has highlighted the scope of change to the DNA methylation landscape in disease states, particularly during tumorigenesis. More recently has been the development of DNA hydroxymethylation profiling techniques, which allows differentiation between 5mC and 5hmC profiles and provides further insights into DNA methylation dynamics and remodeling in tumorigenesis. In this review, we describe the distribution of DNA methylation and DNA hydroxymethylation in different genomic contexts, first in normal cells, and how this is altered in cancer. Finally, we discuss DNA methylation profiling technologies and the most recent advances in single-cell methods, bisulfite-free approaches and ultra-long read sequencing techniques.

## Introduction

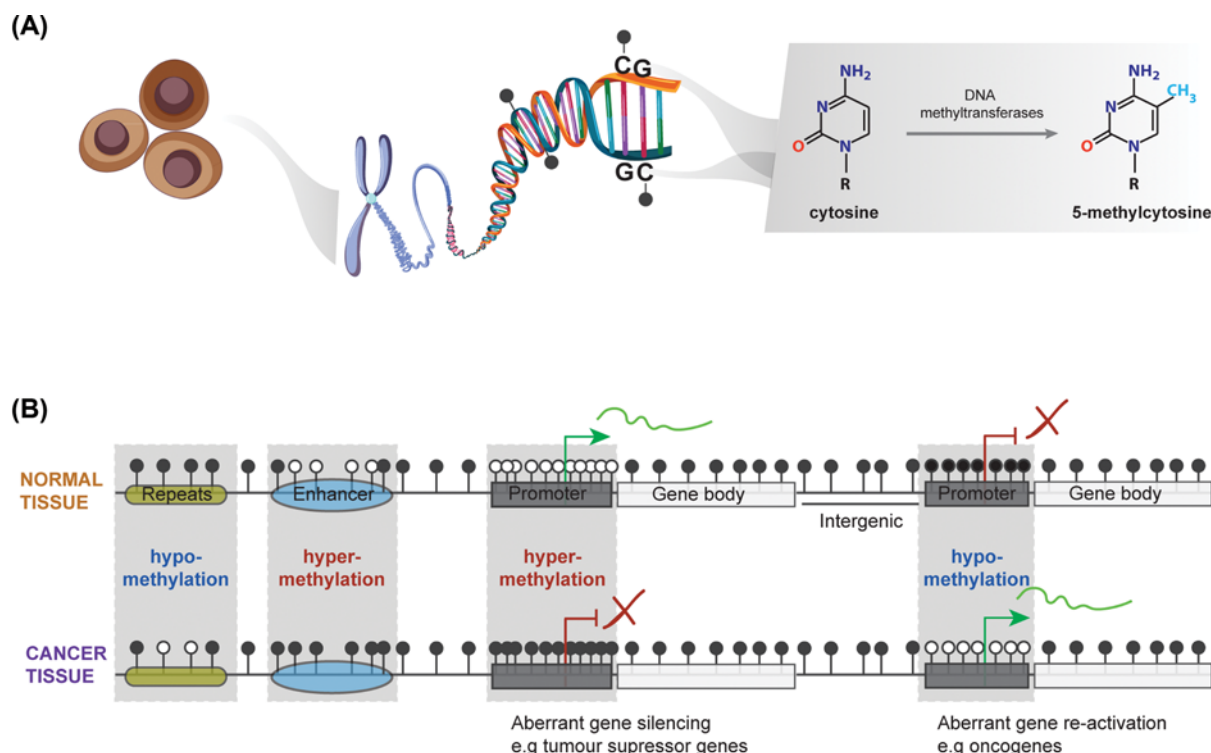
*DNA methylation* in mammalian cells is characterised by the addition of a methyl group at the carbon-5 position of the cytosine base (5-methylcytosine; 5mC) primarily in the context of cytosine-guanine dinucleotides (CpG) through the action of the DNA methyltransferase enzymes (DNMTs) [1] (Figure 1A). Widespread interest in DNA methylation is attributed to the critical role it plays in cell biology [2]; regulating gene expression, retro-element silencing, centromere stability and chromosome segregation in mitosis, X-chromosome inactivation [3,4] and monoallelic silencing of imprinted genes [5].

It is well known that DNA methylation patterns frequently become altered in cancer, including DNA *hypomethylation* events at retro-elements, centromeres and oncogenes in combination with focal DNA *hypermethylation* associated with repression of critical gene regulatory elements such as distal *enhancers* and *promoters* overlapping transcriptional start sites (Figure 1B). Moreover, the discovery that 5mC can be oxidised to *DNA hydroxymethylation* (5-hydroxymethylcytosine; 5hmC) by the ten-eleven translocation (TET) enzymes [6–8] has prompted widespread interest in the possible roles of 5hmC in remodeling the methylation landscape. The ability of TET proteins to further oxidise 5hmC to 5-formylcytosine (5fC) and 5-carboxycytosine (5caC) [9], which can be excised by thymine DNA glycosylase (TDG) in the base excision repair (BER) pathway and replaced with an unmodified cytosine (Figure 2A) provides a mechanism that may contribute to DNA methylation pattern dynamics in the early embryo [10,11], normal cell biology [12,13] and in disease processes [14].

With the advent of genome-wide approaches to interrogate DNA methylation and advances to distinguish major and minor DNA methylation intermediates such as 5hmC from 5mC [15,16], the field is building comprehensive maps of DNA methylation landscapes. Genome-wide mapping studies have revealed that 5hmC and TET proteins are enriched at promoters, gene bodies and distal regulatory elements in mammalian genomes [17,18]. This suggests that the postulated functions of 5mC at these regulatory regions can be revised by taking 5hmC enrichment into account. As we integrate this information from

Received: 19 August 2019  
Revised: 26 November 2019  
Accepted: 26 November 2019

Version of Record published:  
17 December 2019



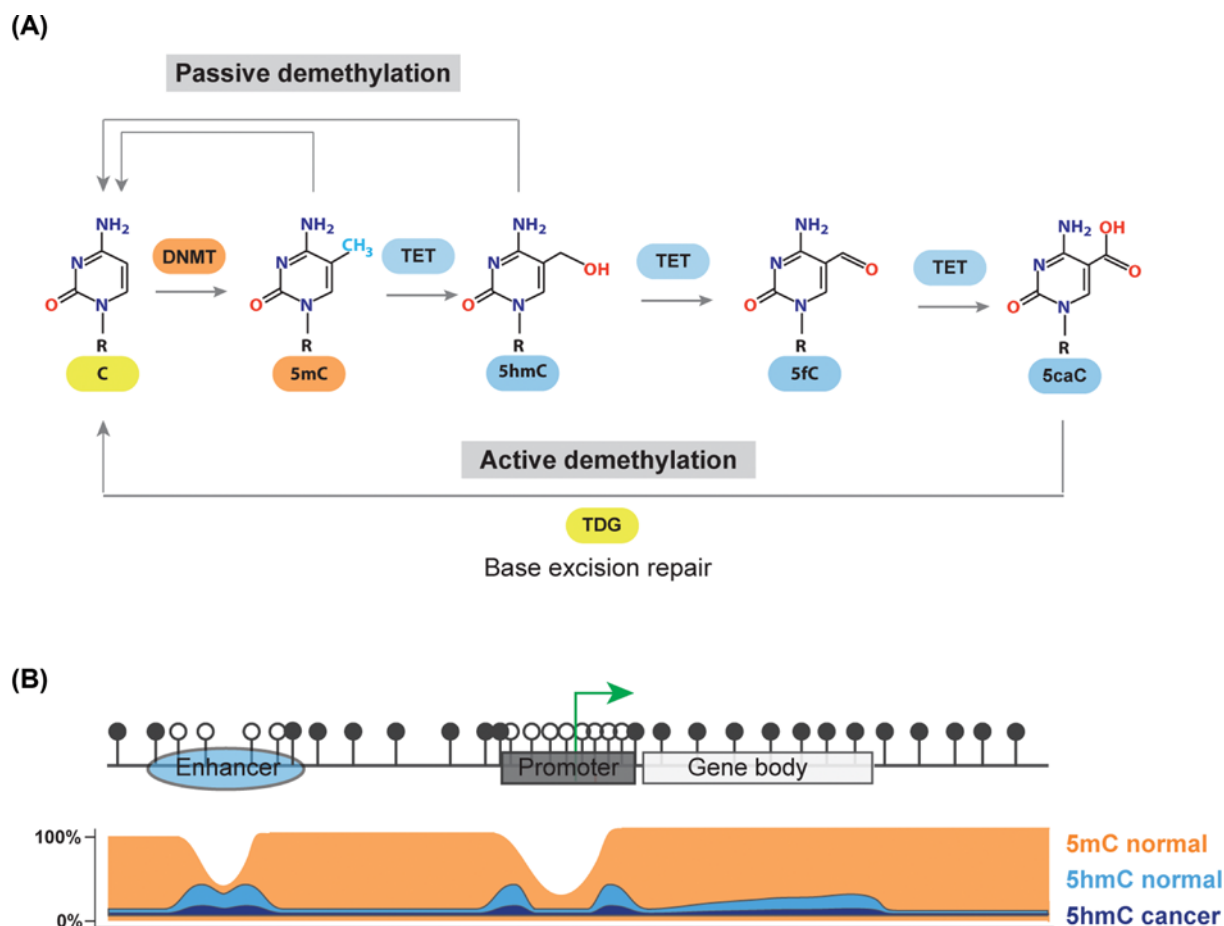
**Figure 1. Normal and cancer genomes exhibit distinct DNA methylation profiles**

(A) Cells utilise DNA methyltransferase (DNMT) enzymes to catalyse the addition of a methyl group to the fifth carbon position of cytosines primarily within CpG dinucleotide contexts (5-methylcytosine; 5mC). This process has various effects on transcription, genome stability and DNA packaging within cells. (B) There are extensive differences in DNA methylation patterning between normal (top) and cancer (bottom) cells across the entire genome, encompassing all gene regulatory elements. The bulk of all CpG sites within the normal genome carry 5mC with distal enhancer elements and CpG island regions being resistant to DNMT activity. The global loss of 5mC is characteristic of cancer cells, with the abnormal presence of punctate increases in DNA methylation across enhancers and promoters. This change in distribution collectively causes a suppression of tumour suppressor genes and concomitant increase in the expression of oncogenes, which drive tumorigenesis. White circle, unmethylated CpG; black circle, methylated CpG. (A) It was made using Vecteezy graphics: Free vector art via <https://www.vecteezy.com>; Vector illustration credit: <https://www.vecteezy.com>.

different cell types and in the context of other epigenetic layers such as *post-translational histone modifications* and nucleosome positions, we are constantly improving our understanding of the role and scope of DNA methylation in different genomic contexts in normal and diseased cells, and importantly, as a function of *tumorigenesis*.

## Normal genomic distribution of DNA methylation

In the human genome, there are approximately 28 million CpG sites of which approximately 70% are methylated in normal somatic cells. Interestingly, these CpG sites are not evenly distributed; in fact, the bulk of the genome is depleted of CpG sites with overall representation of CpG dinucleotides occurring at only one-fifth of the expected frequency. By contrast, clusters of CpG sites occur at the expected frequency, and these regions are termed as ‘CpG islands’ [19]. The majority of CpG islands are 500–1000 base pairs (bp) in length and commonly span promoters of genes, and housekeeping genes [19,20] in particular. Differing from the bulk of the genome, CpG sites located within CpG islands are typically unmethylated in normal somatic cells (Figure 1B). They exist in a transcriptionally permissive chromatin state that is also characterised by various combinations of post-translational histone modifications [21] and distinctive nucleosome organisation [22]. Unmethylated CpG sites within promoter CpG islands provide a binding platform for a complement of transcription factors to control gene activity. A prime example of this is the activity of ubiquitous transcription factor, Sp1 (Specificity protein 1), whose interactions with DNA is modulated by the presence or absence of DNA methylation at CpG islands [23]. Sp1 typically occupies unmethylated DNA to promote gene transcription, whereas binding to methylated CpG sites is inhibited and correlated with transcriptional



**Figure 2. The presence of 5-hydroxymethylation is indicative of both active and passive DNA demethylation pathways**

(A) 5mC formation is implemented by DNMT proteins, and can be sequentially oxidised by TET proteins to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxymethylcytosine (5caC). 5caC can be excised by Thymine DNA glycosylase (TDG) and replaced with an unmodified cytosine through base excision repair (BER). BER-mediated DNA demethylation pathway is referred to as active demethylation. Alternatively, the oxidation of 5mC to 5hmC can compromise DNMT1-mediated maintenance during replication, resulting in passive demethylation of 5mC DNA methylation. (B) Distribution of 5hmC across genomic elements, showing enhancers, promoters and gene bodies. In normal tissue, levels of 5mC are shown in orange and 5hmC is shown in light blue. In cancer, the levels of 5hmC, shown in dark blue, are reduced. White circle, unmethylated CpG; black circle, methylated CpG.

silencing. While the majority of CpG islands are maintained in an unmethylated state, a number of repressed genes harbour methylated promoter CpG islands in somatic cells, which include genes on the inactive X chromosome in females (dosage compensation) and imprinted alleles [24]. Adjacent to CpG islands are regions known as CpG island ‘shores’, which are located approximately two kilobase pairs (2 kB) distal from CpG islands and have comparatively low CpG density [25]. Like CpG island promoters, shores are typically unmethylated in normal cells and this pattern is associated with gene activity.

Gene bodies tend to carry extensive 5mC that, in contrast with CpG island promoters and shores, correlates with active gene expression (Figure 1B). This *intragenic* DNA methylation has been shown to prevent spurious transcription initiation in mouse embryonic stem cells (mESCs) by protecting the gene body from spurious RNA polymerase II binding [26]; furthermore, DNA methylation at CpG islands located *within* gene bodies may repress alternative or tissue-specific promoters [27]. Distal regulatory regions, such as tissue-specific enhancers, are CpG-poor and belong to a class of lowly methylated regions (LMRs) exhibiting average DNA methylation levels ranging between 10 and 50% [28] (Figure 1B). Enhancers possess characteristic post-translational histone modifications [21,29,30] and nucleosome positions [31,32] that work in concert with 5mC to modulate expression of their cognate genes. DNA methylation levels of the enhancer have been shown to be associated with gene activity at validated promoter–enhancer

pairs, with low level of 5mC correlating with increased gene expression [33]. Super-enhancers are regions encompassing clusters of putative enhancers, and here, specific patterns of 5mC have been described ranging from fully methylated to moderately unmethylated and completely unmethylated [34]. Several super-enhancers possess a unique hypermethylated pattern, punctuated by focal unmethylated patches, suggesting non-uniform activity across the super-enhancer. Overall, super-enhancer 5mC levels correlate with the expression of corresponding genes, with completely unmethylated enhancers associated with higher gene activity [34].

## ‘Reshaping’ of 5mC patterns in cancer cells

It is well established that normal epigenetic processes are disrupted during the initiation and progression of tumorigenesis, including global changes to normal DNA methylation patterns [35]. Broadly, this is characterised by overall genome-wide hypomethylation accompanied by regional DNA hypermethylation of CpG island promoters [35–37] (Figure 1B). The hypermethylation of CpG islands is common and frequently associated with silencing of tumour suppressor genes, genes controlling cell growth and downstream signalling pathways. Indeed, numerous loci-specific and genome-wide DNA methylation profiling studies have revealed multiple promoter-associated CpG islands that consistently undergo aberrant DNA hypermethylation in tumour cells. Examples of these include Glutathione S-Transferase P (*GSTP1*) in ~90% of prostate cancers [38], the cyclin-dependent kinase inhibitor, *p16<sup>INK4a</sup>* in ~20% of lung carcinomas [39] and *BRCA1* in ~12% of breast and ovarian carcinomas [40]. It has also been shown that not only are single loci hypermethylated in cancer, but multiple contiguous regions can become coordinately silenced and aberrantly hypermethylated [41]. Furthermore, the frequency of *p16<sup>INK4a</sup>* and *GSTP1* hypermethylation has been shown to increase during disease progression [42,43], suggesting that DNA hypermethylation may be predictive of disease stage or progression. Indeed, The Cancer Genome Atlas consortium first reported the existence of CpG Island Methylator Phenotypes (CIMP) in glioblastoma [44] and colon cancers [45], enabling stratification of disease subtypes by the 5mC signature. Interestingly, the expression of DNMT enzymes is also frequently disrupted in cancer, which provides a feedback loop driving altered DNA methylation patterning across the genome and with potential to cause mutations in the genomic sequence [46]. There is still concerted effort in understanding the similarities between the 5mC profile of embryonic stem cells and cancer cells, including the interesting predisposition of Polycomb marked developmental genes to become preferentially hypermethylated in cancer cells [47–50], and whether the similarities to developmental states represent functional aberrations in the cancer epigenome.

While CpG islands are susceptible to DNA methyltransferase activity, CpG-poor regions tend to undergo hypomethylation during tumorigenesis, resulting in the global decrease in DNA methylation characteristic of tumours [51–53], first described by Feinberg and Vogelstein [54] in colon adenocarcinoma and small cell lung cancer, then later observed in prostate cancer and chronic lymphocytic leukaemia. The exception to this pattern of CpG-poor hypomethylation in cancer cells, is that of CpG-poor enhancer elements; these regions are unmethylated in normal cells and often gain methylation in cancer cells [33,55,56]. In addition, CpG island shores, which also have a lower CpG density, flanking CpG islands up to 2 kb distant, can become DNA hypermethylated in cancer. This was first observed in human colon cancer [25] and also shown in cancer from breast, lung, thyroid and Wilms’ tumour [25,51].

Global DNA hypomethylation in cancer is thought to contribute to genomic instability and increases in aneuploidy [57], both common features of cancer genomes. Indeed, reduced levels of DNA methyltransferase 1 (Dnmt1) can result in increased mutation rate, aneuploidies and tumour development, which provides some evidence for the role of DNA hypomethylation in increased chromosomal fragility [57,58]. Widespread genome instability is commonly accepted to accompany global loss of DNA methylation observed in cancer cells; however causality is still to be definitively shown. Broad regions of global hypomethylation are associated with global changes in chromatin organisation and structural variation [59]. Loss of DNA methylation is also accompanied by aberrant expression of transposable elements, repeat elements and oncogenes [35,57] such as *MYC*, resulting in global deregulation of cellular pathways [60] and occurring concomitant with alterations to chromatin organisation [32] and the three-dimensional genome [56].

There exists a strong link between chromatin structure and DNA methylation, in particular the requirement for a nucleosome to be present for anchoring of DNMT enzymes [61] prior to acquisition of 5mC [62]. Gene expression can be modified in the absence of any changes to DNA methylation patterns, and indeed, developmental gene expression is often already repressed in normal cells but these promoters are prone to abnormal DNA hypermethylation following transformation [50,63]. This is likely due to the fact that the promoters of such developmental genes exhibit a ‘closed’ conformation and well-positioned nucleosomes occlude the transcriptional start site [31]. Organisation of nucleosomes across DNA methylated regions is unexpectedly comparable with genes devoid of 5mC [32] supporting an overall highly organised physical chromatin state irrespective of cell type, normal or cancer origin.



Despite remaining well organised on a broad scale, the substructure of genome organisation is disrupted in cancer cells, further characterised by loss of interactions between enhancer and promoter gene regulatory elements across the cancer genome occurring alongside hotspots of copy number variation [56]. Though a residual level of approximately 5% DNA methylation is required for cell survival [64], the mechanisms driving the initial change in DNA methylation patterning either as a cause or consequence of cancer progression remain to be clearly demonstrated.

## Contribution of 5hmC to the normal DNA methylome

The purpose of 5hmC in the DNA demethylation process has prompted interest in its role in maintaining promoter CpG islands in the unmethylated state in normal cells. In mouse embryonic stem cells (ESCs), 5hmC is depleted from promoters of actively transcribed genes [65,66] but enriched at bivalent promoters of poised genes in a bimodal fashion surrounding the transcription start sites (Figure 2B) [17,67]. Similarly, in non-pluripotent cells, 5hmC demarcates the borders of promoter CpG islands [68–70] and large undermethylated regions called CpG canyons [71]. Levels of 5hmC accumulate at gene bodies during neuronal differentiation [72,73] and positively correlate with gene expression [13,73–76] (Figure 2B). The correlation between 5hmC and gene activity is more pronounced than that of 5mC [77]. Notably, 5hmC accumulation at gene bodies during differentiation was not accompanied by subsequent DNA demethylation, suggesting that 5hmC is a stable epigenetic mark at gene bodies in the brain [72]. Such a role would imply the existence of 5hmC binding proteins [75,78] and/or the ability of 5hmC to repel the recruitment and counteract the activity of 5mC binding proteins. This is supported by the observation that a subset of the methyl-CpG binding domain (MBD) family of proteins, while possessing high affinity to 5mC, were unable to bind 5hmC containing DNA [79]. There has also been interest in the role of intragenic 5mC [80] and 5hmC [76,81,82] in the regulation of alternative splicing, particularly at exon–intron boundaries in brain tissue. The presence of 5mC, together with 5hmC and specific histone modifications [83–85], at gene bodies may establish specific chromatin signatures and nucleosome positioning within the transcribed unit to regulate RNA transcription kinetics, gene activity and splicing. However, the role for 5hmC in shaping the regulatory function of these proteins is yet to be elucidated. Importantly, TET triple knockout in mouse and human ESCs results in increased DNA methylation at the borders of bivalent but not active, promoters [6,86]. Similarly, in differentiated cells, TET1 depletion triggers DNA methylation spreading into promoter CpG islands from the methylated borders [68] highlighting the crucial role for TET proteins and 5hmC in the maintenance of DNA methylation boundaries. Depletion of DNMT 3A isoform 1 (DNMT3A1), in turn, results in erosion of 5mC concomitant with 5hmC reduction at the shores of bivalent CpG islands [87] and CpG canyons [71].

Much like the accumulation of 5hmC at the borders of bivalent promoters, 5hmC is normally highly enriched at distal regulatory elements (Figure 2B) coinciding with enhancer-associated post-translational histone marks H3K4me1 and H3K27ac [88], DNase hypersensitive sites [89] and LMRs [90]. 5hmC is enriched at enhancers in a bimodal fashion surrounding the TF/p300 binding site [89] as well as at the boundaries of super-enhancers [91], coinciding with a local depletion of 5mC [76,89,91], which increases in response to TET depletion [6,92]. Moreover, in TDG-depleted cells, the majority of induced accumulation of 5caC occurs at poised and active enhancers [66] supporting the idea that TET-mediated 5mC oxidation followed by TDG-mediated excision is essential for the retention of reduced 5mC levels at distal regulatory elements. In fact, TET2 depletion-mediated hypermethylation predominantly occurs at distal regulatory elements to the highest extent at ‘weak’ enhancers with lower enrichment of H3K27ac and lower TF occupancy, but higher baseline levels of 5mC methylation [92]. Therefore, weak enhancers might be subject to a more dynamic interplay between DNMT-mediated methylation and TET-mediated demethylation, potentially governing enhancer activity through the mediation of TF binding. Altogether, these observations suggest that the balance between DNA methylation and DNA demethylation is required for the maintenance of methylation landscape at promoter borders. Consequently, an impairment of this balance may result in DNA methylation spreading into the interior of the promoter CpG island or, oppositely, erosion of DNA methylation boundary and formation of hypomethylated CpG island shores with both scenarios shown as a feature of carcinogenesis [51]. The cross-talk between 5mC, 5hmC and chromatin modifications at enhancer elements is complex and further research into the role of DNA methylation at distal regulatory elements is ongoing [93].

## Rearrangement of the 5hmC landscape in cancer cells

Alongside the global loss of 5mC, a widespread reduction in 5hmC has been observed in all studied human carcinomas of different histological types [94,95]. Dot blot hybridisation and immunohistochemistry analyses have revealed that levels of 5hmC are significantly reduced in melanoma [94], breast, prostate, colon [95], liver, lung and pancreatic cancer [96] compared with the adjacent normal tissues. Quantitative estimates of 5hmC reduction in cancer have

confirmed significant depletion of 5hmC up to 5-fold in small cell lung carcinomas and 30-fold in astrocytomas and glioblastomas [97]. It is likely that reduced levels of 5hmC observed in cancer cells are due to impaired activity of TET enzymes or a reduction in expression in TET proteins [94,96]. A number of cancers also carry mutations in TET enzymes with a high recurrence of TET2 mutations in myeloid malignancies [98,99]. Tumours carrying TET mutations display reduced 5hmC [100,101], and TET depletion reduces 5hmC levels in human leukaemia cell lines [102,103]. In the absence of TET aberrations, an indirect effect can be observed via altered enzymatic activity of Isocitrate Dehydrogenases 1 and 2 (IDH1/2) [104]. IDH1/2 catalyses the oxidative decarboxylation of isocitrate to 2-oxoglutarate, which serves as an essential cofactor of TET catalytic activity [105]. IDH1/2 mutations result in the reduction of 2-oxoglutarate and aberrant accumulation of (R)-2-hydroxyglutarate, a competitive inhibitor of TET catalytic activity [106–108] were found in melanoma [94, glioma [44,109] and acute myeloid leukaemia [110]. Consistent with TET mutations, tumours carrying IDH mutations displayed increased CpG island hypermethylation [44,106]. The reintroduction of catalytically active TET2 or IDH2 results in restoration of the 5hmC landscape and suppression of melanoma invasion and growth [94]. Similar effects are seen following the reintroduction of catalytically active TET1 in normal and tumour breast cells [111]. These data highlight a vital role for TET and IDH proteins in the maintenance of the DNA methylation and DNA hydroxymethylation landscape in normal cells and their alteration in tumorigenesis. However, the function of TET or IDH enzymes is impaired only in a minority of cancers, whereas widespread 5hmC reduction has been shown to occur in all cancer tissues studied to date [95–97]. Thus, this suggests that a more generalised mechanism exists for 5hmC loss in cancer.

Critically, apart from IDH/2-oxoglutarate, the activity of TET as well as many other chromatin-modifying enzymes is dependent on several other cofactors such as  $\text{Fe}^{2+}$ , oxygen and ascorbic acid. Oxygen availability is a crucial regulator of TET catalytic activity. It enables 5mC to 5hmC oxidation by oxidising  $\text{Fe}^{2+}$  in TET catalytic pocket and inducing oxidative decarboxylation of 2-oxoglutarate. Indeed, 5hmC loss has been attributed to tumour oxygenation with hypoxic areas within patient-derived tumour xenografts exhibiting decreased 5hmC [112]. Ascorbic acid is another essential cofactor of TET enzymes serving as an electron donor in the  $\text{Fe}^{3+}$  to  $\text{Fe}^{2+}$  reduction reaction. Supplementation of ascorbic acid was shown to enhance TET catalytic activity and transient 5hmC accumulation in mouse ESCs [113] and embryonic fibroblasts [114]. In melanoma and bladder cancer cells ascorbic acid enabled restoration of 5hmC levels and this was sufficient to inhibit cancer cell growth and migration [115,116]. Overall, TET/IDH mutations, hypoxia and ascorbic acid depletion can attenuate the catalytic function of TET enzymes potentially serving as underlying cause of widespread 5hmC loss in cancer.

## DNA methylation profiling technologies

Outstanding biological questions remain regarding the DNA methylation process itself, as well as how genes are silenced and targeted for silencing. In addition, the field is at an inflection point where it must adopt new techniques that circumvent current limitations such as cost, genomic coverage and sample input requirements so that the study of 5mC and 5hmC become standard approaches in the clinic for diagnostics and monitoring of tumorigenesis.

Technologies to measure DNA methylation are long-established but constantly evolving. In particular, the potential to examine the *genome-wide* presence or absence of 5mC using next-generation sequencing techniques has also enabled assessment of DNA methylation at the level of individual CpG dinucleotides. By contrast, some of the earliest techniques such as high performance liquid chromatography (HPLC, [117]) measure *global* 5mC content. As noted by Vryer and Saffery [118], the distinction is important, as direct approaches remain the only current techniques that determine global methylation by definition. This is affirmed by the presence of persistent ‘gaps’ in genomic sequencing coverage meaning that the majority (approximately 95%), but not all, CpG sites can be analysed using the whole-genome bisulfite sequencing technique [119]. While we are in an excellent position to perform comprehensive characterisation of DNA hyper- and hypo-methylation events in distinct tumour types, the field has simultaneously exploded into the single cell space and is now developing bisulfite-free DNA methylation assays to circumvent some of the existing limitations in measuring DNA methylation. The overarching goal in cancer biology is to apply these cutting-edge advances to detect aberrant 5mC and 5hmC changes to gain insights into common versus distinct regulatory pathways. Perhaps more importantly, to provide a position from which to determine the prerequisites underlying the aberrant 5mC and 5hmC changes, and new options for manipulating them for the benefit of patients.

## Global technologies

Global measurement of DNA methylation defines the total amount of 5-methylcytosine relative to unmethylated cytosine content in the genome. This is *directly* achieved using HPLC [117], high-performance capillary electrophoresis

(HPCE, [120]) or liquid chromatography in combination with tandem mass spectrometry (LC-MC/MS, [121]). Generally, direct measurement approaches are not commonly used due to the need for relatively large amounts of sample DNA, specialist equipment and being unsuitable for high-throughput processing.

Indirect measurement of global DNA methylation levels is achieved using any other method. The overarching assumption of indirect measurements is that the average methylation level of the subsampled CpG dinucleotides is representative (proxy) of the remainder of the genome. This category of assays includes subsampling of repeat genomic elements, enzymatic digestion with or without dependence on the presence of methyl groups at consensus sequences containing CpG sites (for example, by using *HpaII* or *MspI* enzymes) and luminometric methylation (LUMA, [122]) or enzyme-linked immunosorbent assays (ELISA). A proxy measurement is most accurate using the long interspersed numerical element (LINE) assay when compared with HPLC [123] and most variable using standard ELISA-based approaches (for example, the 5mC DNA ELISA Kit from Zymo Research). However, the specific biological question, sample type, budget and available equipment are all important considerations when deciding on a method to assess DNA methylation. Thus, ELISA-based approaches that use specific antibodies to detect 5mC or 5hmC are appropriate for high throughput screening or when the goal is to identify large changes (>1.5-fold) between samples.

## Genome-wide technologies

Next-generation sequencing platforms have revolutionised biology, notably for assessment and interpretation of DNA methylation events. Foremost, the ability to achieve base pair resolution and assess methylation outside dense, *CpG island* regions usually found at gene promoters [19,20]. Whole-genome bisulfite sequencing (WGBS) was first demonstrated by Lister et al. [124] by adaptation of the current gold-standard from Frommer and Clark [125]. This approach is becoming more accessible as a consequence of improvements in workflow (for example, picogram amounts of DNA input) and reduced costs of genomic sequencing; yet, <2000 WGBS datasets have been deposited into the Gene Expression Omnibus (GEO) database to date. This reflects the relative expense of a WGBS dataset compared with other options, requirement for technical expertise and significant data analysis. For these reasons, WGBS is not yet routinely used in research or clinical laboratories. Nucleosome Occupancy and Methylation Sequencing (NOME-Seq [32,126]) takes advantage of the WGBS protocol, but offers improvement by simultaneously measuring endogenous nucleosome occupancy and chromatin accessibility. This is achieved by treating nuclei with bacterial *M.CviPI* enzyme, which gives an exogenous methylation profile of the accessible GpC nucleotides (nucleosomes and accessibility) without affecting endogenous CpG methylation. Samples are otherwise prepared, and cost is equivalent to WGBS. Enrichment-based methods such as MeDIP [127] and MBDCap-Seq [128] tend to produce comparatively 'noisy' data and report on an ~18% fraction of CpG sites. Moreover, MeDIP favours regions of low CpG density such as intergenic regions while MBDCap-Seq is biased towards CpG-rich elements. The cost to sequence an enriched sample is at least 15-fold less than WGBS or NOME-Seq, but reduced bias can be achieved with more recent array-based assays for similar expense. Indeed, the latest iteration of the Illumina Infinium Array (MethylationEPIC) remains one of the most popular selections for DNA methylation analysis, particularly for clinical samples. Data from the ~850 000 CpG dinucleotides covered by the MethylationEPIC reliably mirrors WGBS measurements of identical sites [129] and covers ~3% of all CpG dinucleotides located within a spectrum of genomic regulatory elements including enhancers and promoters. While biological interpretation of the MethylationEPIC data rely on the 'co-methylation assumption' of adjacent CpG sites [130], there is an increasing number of user-friendly data analysis options that improve accessibility of array-based methods (e.g. CHAMP [131]). Targeted bisulfite sequencing offers similar benefits with increased coverage and the potential for customisation (for example, CpGiant Enrichment System (Roche) can assess up to 5.5 million CpG sites) but has a greater sample input requirement. The BLUEPRINT consortium [132] evaluated performance, sensitivity, scalability and cost of 27 locus-specific and genome-wide DNA methylation assays for clinical applications. This report highlighted the relative strengths and weaknesses of common assays; of the current methods available, the consortium recommended locus-specific amplicon bisulfite sequencing and pyrosequencing methods for overall performance, noting that currently popular genome-wide alternatives (such as MethylationEPIC) have comparatively reduced accuracy and higher cost.

## Single-cell advances, bisulfite-free and ultra-long read sequencing approaches

Existing next-generation sequencing based assays usually use a population of cells to achieve the input requirement and typically do not distinguish 5mC from 5hmC unless modified protocols are employed (for example, TET-assisted bisulfite sequencing [89]). Thus, the field has been adapting and developing techniques that allow interrogation of cell-type specific methylation, in single cells and/or independently of a bisulfite conversion step.

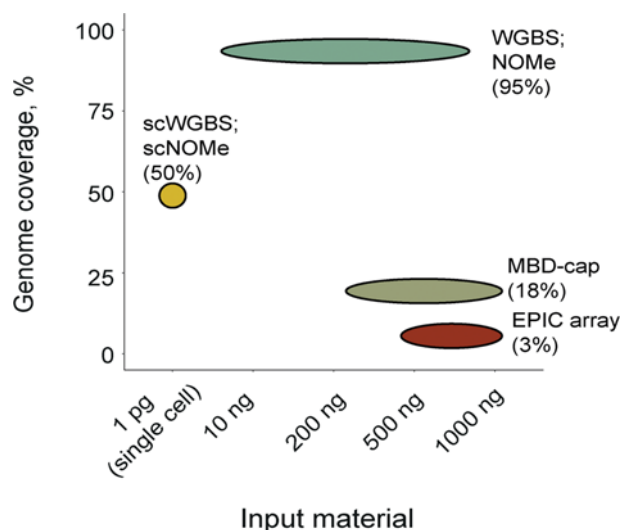
The development of single cell bisulfite sequencing [133] (scBS-Seq) was a breakthrough being the first epigenomic methodology adapted for single cells. However, the initial datasets produced in oocytes and embryonic stem cells covered fewer than half of all possible CpG sites (up to ~48%) in the mouse genome and data mapping efficiency was <25% due to low sample complexity [133]. Variations on scBS-Seq have seen the ability to measure DNA methylation and transcription in parallel from single cells (scM&T-Seq, [134]) and single cell nucleosome occupancy and methylation sequencing (scNOME-Seq, [135]). Wang et al. [136] have recently devised a new approach (methyltransferase treatment followed by single-molecule long-read sequencing (MeSMLR-Seq) based upon the original NOME-seq protocol, which adopts single molecule long read sequencing (Oxford Nanopore Technologies) to enable in excess of 53 kB read lengths (up to 356 nucleosomes). MeSMLR-Seq offers one of the first epigenetic applications of long-read sequencing and great potential to uncover biological insight over short-read sequencing options; however, there is currently little cost or convenience benefit in adopting these approaches in their current form.

The vast majority of current DNA methylation assays rely on bisulfite conversion, including interrogation of 5hmC. This is problematic because the process of bisulfite conversion itself damages 84–99% of the genomic DNA [137,138]. While single cell and low input options exist for bisulfite-dependent methods including WGBS, there is a need to further develop bisulfite-independent options. ACE-Seq takes advantage of human-specific APOBEC3A enzyme to resolve 5hmC distribution without bisulfite conversion [139] and chemical labelling methods increase sensitivity and specificity of the 5hmC readout at single molecule resolution [74,140]. Adaptations of chemical labelling approaches based on 5fC-T transition have enabled mapping of the low abundance 5fC intermediate in bulk cells [141,142] and at a single cell level [143]. Using three variations of a pic-borane treatment coupled with exogenous TET enzyme treatment, Liu et al. [144] have resolved 5mC (TET-assisted pyridine borane sequencing  $\beta$ ; TAPS $\beta$ ), 5hmC (chemical-assisted borane sequencing; CAPS) and 5mC+5hmC (TET-assisted pyridine borane sequencing; TAPS). These methods preserve double-stranded DNA without excessive fragmentation, resulting in greater coverage of CpG sites without sequence bias [144]. These methods are still in their relative infancy and their uptake is not yet widespread; however, we predict that it will be the future adaptation of these bisulfite-free assays in a clinical setting, likely in combination with the genome-wide sequencing applications currently in use, that will be groundbreaking for patient treatments or refined diagnoses. In part, because of the information retention possible without damaging bisulfite treatment but additionally, because of the reduced amounts of DNA sample required for input. The next advances will also require increased capability to map the three-dimensional genome at an extraordinarily fine resolution, alongside the methylome, which is already achieved at base-pair resolution. As with all massive datasets, the difficulty extends beyond the ability to produce the data in the first place, and the field will be faced with ongoing challenges regarding the interpretation of methylomes either standalone or in combination with genetic, transcriptomic and/or complementary epigenetic information (e.g. post-translational histone modification screening) for patient benefit.

## Concluding remarks

We are at an inflection point in generating and interpreting cancer epigenomes, particularly DNA methylation. With comprehensive mapping of cancer-associated DNA methylation changes now possible, we are faced with the necessity to improve the cost: accessibility and genome coverage: input ratios of genome-wide approaches (Figure 3). This is particularly important as bisulfite-free methods with base pair resolution, single molecule and single cell capability become possible. There is need for excitement around these approaches because they address three pitfalls of whole-genome bisulfite sequencing; namely, (1) the input requirements are compatible with a wider range of sample types, (2) DNA integrity is preserved and (3) ‘CpG dropout’ or ‘gaps’ are avoided. The ultimate benefit is greater or complete coverage of the 28 million CpG sites is achieved for less cost. However, from a biological perspective there are many open questions that remain despite great advances in the past decade. The mechanisms underlying DNA methylation changes incorporating global hypomethylation alongside punctate DNA hypermethylation events are still to be characterised. Why are CpG islands, shores and distal enhancer elements susceptible to DNA methyltransferase activity in cancer? Indeed, which of these are driving cancer initiation and progression? How dynamic is a change in DNA methylation at any given locus or CpG site and does fluctuate across disease trajectories? The ability to profile single cells will not offer an explanation to this end, and carefully designed animal and cell line experiments will be required to monitor DNA methylation and DNA hydroxymethylation flux. These platforms, in addition to being able to monitor DNA methylomes in clinical samples, are a critical base to extend our understanding of DNA methylation biology in normal and cancer cell contexts.





**Figure 3. Comparison of genome coverage in input requirements of common assays to measure DNA methylation**

Technologies to measure DNA methylation (5mC) typically rely on bisulfite conversion of cytosine into uracil, while 5mC is resistant to conversion. Adaptations (for example, TET-assisted bisulfite sequencing, allow the measurement of 5hmC. Nucleosome occupancy and methylation sequencing (NOME) and WGBS are amenable to various amounts of input (~10–1000 ng genomic DNA) resulting in up to 95% coverage of CpG sites across the genome. The single cell variants of NOME and WGBS both result in 50% loss of coverage. The Methylation EPIC array still requires substantial DNA input (~500–1000 ng) and covers only 3% of all CpG sites, but is particularly favourable for clinical samples. MBDCap-Seq requires input of ~0.2–1000 ng genomic DNA and covers approximately 18% of the genome.

## Summary

- DNA methylation changes are widespread in cancer cells, including both DNA hyper- and hypomethylation events and DNA hydroxymethylation alterations.
- DNA methylation and DNA hydroxymethylation changes occur across the genome with emphasis on changes occurring distal from promoters and the need to understand the broader impact of these changes on gene expression, genome structure and cell behaviour.
- Many options are available to assess DNA methylation on global, locus-specific and genome-wide scales in cell populations as well as single cells.
- The field is developing new methods to circumvent the disadvantages of current techniques.
- A focus on data interpretation is also important, as the number of available datasets is continually increasing.

## Author Contribution

K.S., C.S. and P.T. contributed equally to writing of this manuscript.

## Funding

This work was supported by the National Health and Medical Research Council (NHMRC) Project Grants [grant numbers APP 1128912 (to C.S.), APP 1161985 (to P.T.), APP 1109696 (to P.T.)]; the National Breast Cancer Foundation Investigator Initiated Research Scheme Grant [grant number IIRS-18-137 (to C.S.)]; the National Foundation and Medical Research and Innovation (NFMRI) grant [grant number NFMRI\_Stirzaker (to C.S.)]; and Cancer Council Tasmania (Australia) Grants Scheme [grant number T24716 (to P.T.)].

## Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

## Glossary

5-methylcytosine, 5mC; methylated form of the fifth carbon position of the cytosine nucleotide in deoxyribonucleic acid (DNA); 5-hydroxymethylcytosine, 5hmC; DNA pyrimidine nitrogen base derived as the first oxidative product in the process of demethylating cytosine nucleotides. Oxidation is facilitated by Ten-Eleven-Translocation (TET) enzymes. Precursor to 5-formylcytosine (5fC), then 5-carboxylcytosine (5caC); Bisulfite sequencing, shotgun or genome-wide sequencing methods applied to bisulfite-treated DNA allowing measurement of cytosine methylation at single base pair resolution. Currently considered the 'gold standard', it is theoretically but not practically possible to resolve every cytosine using bisulfite sequencing, which also does not distinguish between 5mC and 5hmC; Coverage, 'depth'; referring to the number sequencing reads covering any given nucleotide in the genome. The average coverage is typically reported for sequencing datasets; CpG island, defined as a genomic region greater than 500 bp with a G+C content equal to or greater than 55% and observed CpG/expected CpG of 0.65 [20]. This definition identified islands more likely to be associated with the 5' regions of genes and excluded the majority of Alu-repetitive elements; Cytosine, C; one of four nucleotide bases found in DNA and ribonucleic acid (RNA); Direct, referring to the absolute measurement of 5mC in a sample (e.g. by HPLC) and distinct from most genome-wide methods available that measure a subset of all CpG sites and extrapolate the total percent methylation (i.e. indirectly measuring global DNA methylation levels); DNA methylation, Usually refers to the presence of 5mC and the activity of DNA methyltransferase (DNMT) enzymes. However, the term 'DNA methylation' could encompass DNA hydroxymethylcytosine and other variants, as well as methylation in non-CpG contexts; DNA hydroxymethylation, Referring to the presence of 5hmC; ELISA, enzyme-linked immunosorbent assay (commercially available; MethylFlash Methylated DNA 5mC Quantification Kit (Epigentek) or EpiSeeker (Abcam) for example; Enhancer, A regulatory region DNA of variable (50–1000 bp) size capable of binding transcriptional and chromatin regulatory proteins for the purpose of controlling gene expression. Often located distal to the promoter, they usually interact with their cognate genes by DNA looping; Genome-wide, 'across', referring to measurement of 5mC at a representative subset (1.7–95% using current assays) of all cytosine residues across the genome; Global, 'total', referring to the total level of 5-methylcytosine relative to unmethylated cytosine content; Hypomethylation, Loss/reduction in methylation compared with normal or what is expected (e.g. hypomethylation of the bulk of the genome in cancer); Hypermethylation, Gain/increase in methylation compared with normal or what is expected (e.g. hypermethylation of CpG islands in cancer); Indirect, referring to the assumption that 5mC levels at a subset or majority of CpG sites accurately reflects the global DNA methylation content; Intergenic, located between or outside a gene/s; Intragenic, located within a gene/s; LUMA, Luminometric methylation assay; couples restriction enzyme digestion with pyrosequencing; Nucleosome, an octamer of eight histone proteins; facilitates the compaction of the DNA strand inside the nucleus, influences DNA accessibility profiles and provides a platform for the binding of enzymes that post-translationally modify the histone proteins; Post-translational histone modification, Covalent modification of histone proteins following translation and including the processes of acetylation, phosphorylation and methylation, for example; Promoter, Located proximal (5'-) to the transcriptional start site, promoters are 500–1000 bp and can contain many defined regulatory elements. Promoters are the sites of binding for many transcription factors and chromatin modifying proteins required for transcription (e.g. RNA Polymerase II); 'Shores', Immediately flanking and up to ~2 kb distal from CpG islands; 'shores' are regions of relatively low CpG density compared with CpG islands. The presence of DNA methylation at 'shores' is correlated with gene silencing; Tumorigenesis, formation and evolution of a tumor.

## References

- Bestor, T.H. (1988) Cloning of a mammalian DNA methyltransferase. *Gene* **74**, 9–12, [https://doi.org/10.1016/0378-1119\(88\)90238-7](https://doi.org/10.1016/0378-1119(88)90238-7)
- Bird, A.P. (1986) CpG-rich islands and the function of DNA methylation. *Nature* **321**, 209–213, <https://doi.org/10.1038/321209a0>
- Mohandas, T., Sparkes, R.S. and Shapiro, L.J. (1981) Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science* **211**, 393–396, <https://doi.org/10.1126/science.6164095>
- Gartler, S.M. and Riggs, A.D. (1983) Mammalian X-chromosome inactivation. *Annu. Rev. Genet.* **17**, 155–190, <https://doi.org/10.1146/annurev.ge.17.120183.001103>
- Reik, W., Collick, A., Norris, M.L., Barton, S.C. and Surani, M.A. (1987) Genomic imprinting determines methylation of parental alleles in transgenic mice. *Nature* **328**, 248–251, <https://doi.org/10.1038/328248a0>
- Lu, F., Liu, Y., Jiang, L., Yamaguchi, S. and Zhang, Y. (2014) Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev.* **28**, 2103–2119, <https://doi.org/10.1101/gad.248005.114>
- Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A. et al. (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473**, 398–402, <https://doi.org/10.1038/nature10008>
- Bogdanovic, O., Smits, A.H., de la Calle Mustienes, E., Tena, J.J., Ford, E., Williams, R. et al. (2016) Active DNA demethylation at enhancers during the vertebrate phylotypic period. *Nat. Genet.* **48**, 417–426, <https://doi.org/10.1038/ng.3522>
- Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A. et al. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–1303, <https://doi.org/10.1126/science.1210597>

- 10 Shen, L., Inoue, A., He, J., Liu, Y., Lu, F. and Zhang, Y. (2014) Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. *Cell Stem Cell* **15**, 459–471, <https://doi.org/10.1016/j.stem.2014.09.002>
- 11 Guo, F., Li, X., Liang, D., Li, T., Zhu, P., Guo, H. et al. (2014) Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell* **15**, 447–459, <https://doi.org/10.1016/j.stem.2014.08.003>
- 12 Wang, T., Pan, Q., Lin, L., Szulwach, K.E., Song, C.X., He, C. et al. (2012) Genome-wide DNA hydroxymethylation changes are associated with neurodevelopmental genes in the developing human cerebellum. *Hum. Mol. Genet.* **21**, 5500–5510, <https://doi.org/10.1093/hmg/dds394>
- 13 Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D. et al. (2013) Global Epigenomic Reconfiguration During Mammalian Brain Development. *Science* **341**, 1237905, <https://doi.org/10.1126/science.1237905>
- 14 Ficiz, G. and Gribben, J.G. (2014) Loss of 5-hydroxymethylcytosine in cancer: cause or consequence? *Genomics* **104**, 352–357, <https://doi.org/10.1016/j.ygeno.2014.08.017>
- 15 Booth, M.J., Branco, M.R., Ficiz, G., Oxley, D., Krueger, F., Reik, W. et al. (2012) Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science* **336**, 934–937, <https://doi.org/10.1126/science.1220671>
- 16 Yu, M., Hon, G.C., Szulwach, K.E., Song, C.X., Jin, P., Ren, B. et al. (2012) Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. *Nat. Protoc.* **7**, 2159–2170, <https://doi.org/10.1038/nprot.2012.137>
- 17 Wu, H., D'Alessio, A.C., Ito, S., Wang, Z., Cui, K., Zhao, K. et al. (2011) Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev.* **25**, 679–684, <https://doi.org/10.1101/gad.203601>
- 18 Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J. et al. (2011) Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol. Cell* **42**, 451–464, <https://doi.org/10.1016/j.molcel.2011.04.005>
- 19 Gardiner-Garden, M. and Frommer, M. (1987) CpG islands in vertebrate genomes. *J. Mol. Biol.* **196**, 261–282, [https://doi.org/10.1016/0022-2836\(87\)90689-9](https://doi.org/10.1016/0022-2836(87)90689-9)
- 20 Takai, D. and Jones, P.A. (2002) Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 3740–3745, <https://doi.org/10.1073/pnas.052410099>
- 21 Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D. et al. (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39**, 311–318, <https://doi.org/10.1038/ng1966>
- 22 Lai, W.K.M. and Pugh, B.F. (2017) Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat. Rev. Mol. Cell Biol.* **18**, 548–562, <https://doi.org/10.1038/nrm.2017.47>
- 23 Butta, N., Larrucea, S., Alonso, S., Rodriguez, R.B., Arias-Salgado, E.G., Ayuso, M.S. et al. (2006) Role of transcription factor Sp1 and CpG methylation on the regulation of the human podocalyxin gene promoter. *BMC Mol. Biol.* **7**, 17, <https://doi.org/10.1186/1471-2199-7-17>
- 24 Li, E., Beard, C. and Jaenisch, R. (1993) Role for DNA methylation in genomic imprinting. *Nature* **366**, 362–365, <https://doi.org/10.1038/366362a0>
- 25 Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P. et al. (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat. Genet.* **41**, 178–186, <https://doi.org/10.1038/ng.298>
- 26 Neri, F., Rapelli, S., Krepelova, A., Incarnato, D., Parlato, C., Basile, G. et al. (2017) Intragenic DNA methylation prevents spurious transcription initiation. *Nature* **543**, 72–77, <https://doi.org/10.1038/nature21373>
- 27 Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D'Souza, C., Fouse, S.D. et al. (2010) Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466**, 253–257, <https://doi.org/10.1038/nature09165>
- 28 Stadler, M.B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Scholer, A. et al. (2011) DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490–495, <https://doi.org/10.1038/nature10716>
- 29 Creighton, M.P., Cheng, A.W., Weststead, G.G., Kooistra, T., Carey, B.W., Steine, E.J. et al. (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 21931–21936, <https://doi.org/10.1073/pnas.1016071107>
- 30 Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A. and Wysocka, J. (2011) A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283, <https://doi.org/10.1038/nature09692>
- 31 Taberlay, P.C., Kelly, T.K., Liu, C.C., You, J.S., De Carvalho, D.D., Miranda, T.B. et al. (2011) Polycomb-repressed genes have permissive enhancers that initiate reprogramming. *Cell* **147**, 1283–1294, <https://doi.org/10.1016/j.cell.2011.10.040>
- 32 Taberlay, P.C., Statham, A.L., Kelly, T.K., Clark, S.J. and Jones, P.A. (2014) Reconfiguration of nucleosome-depleted regions at distal regulatory elements accompanies DNA methylation of enhancers and insulators in cancer. *Genome Res.* **24**, 1421–1432, <https://doi.org/10.1101/gr.163485.113>
- 33 Aran, D., Sabato, S. and Hellman, A. (2013) DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. *Genome Biol.* **14**, R21, <https://doi.org/10.1186/gb-2013-14-3-r21>
- 34 Heyn, H., Vidal, E., Ferreira, H.J., Vizoso, M., Sayols, S., Gomez, A. et al. (2016) Epigenomic analysis detects aberrant super-enhancer DNA methylation in human cancer. *Genome Biol.* **17**, 11, <https://doi.org/10.1186/s13059-016-0879-2>
- 35 Jones, P.A. and Baylin, S.B. (2007) The epigenomics of cancer. *Cell* **128**, 683–692, <https://doi.org/10.1016/j.cell.2007.01.029>
- 36 Baylin, S.B. and Jones, P.A. (2011) A decade of exploring the cancer epigenome - biological and translational implications. *Nat. Rev. Cancer* **11**, 726–734, <https://doi.org/10.1038/nrc3130>
- 37 Jones, P.A. and Baylin, S.B. (2002) The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* **3**, 415–428, <https://doi.org/10.1038/nrg816>
- 38 Millar, D.S., Ow, K.K., Paul, C.L., Russell, P.J., Molloy, P.L. and Clark, S.J. (1999) Detailed methylation analysis of the glutathione S-transferase pi (GSTP1) gene in prostate cancer. *Oncogene* **18**, 1313–1324, <https://doi.org/10.1038/sj.onc.1202415>
- 39 Herman, J.G., Merlo, A., Mao, L., Lapidus, R.G., Issa, J.P., Davidson, N.E. et al. (1995) Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers. *Cancer Res.* **55**, 4525–4530
- 40 Esteller, M., Silva, J.M., Dominguez, G., Bonilla, F., Matias-Guiu, X., Lerma, E. et al. (2000) Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. *J. Natl. Cancer Inst.* **92**, 564–569, <https://doi.org/10.1093/jnci/92.7.564>

- 41 Coolen, M.W., Stirzaker, C., Song, J.Z., Statham, A.L., Kassir, Z., Moreno, C.S. et al. (2010) Consolidation of the cancer genome into domains of repressive chromatin by long-range epigenetic silencing (LRES) reduces transcriptional plasticity. *Nat. Cell Biol.* **12**, 235–246, <https://doi.org/10.1038/ncb2023>
- 42 Mahon, K.L., Qu, W., Devaney, J., Paul, C., Castillo, L., Wykes, R.J. et al. (2014) Methylated Glutathione S-transferase 1 (mGSTP1) is a potential plasma free DNA epigenetic marker of prognosis and response to chemotherapy in castrate-resistant prostate cancer. *Br. J. Cancer* **111**, 1802–1809, <https://doi.org/10.1038/bjc.2014.463>
- 43 Belinsky, S.A., Nikula, K.J., Palmisano, W.A., Michels, R., Saccomanno, G., Gabrielson, E. et al. (1998) Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 11891–11896, <https://doi.org/10.1073/pnas.95.20.11891>
- 44 Noshmeh, H., Weisenberger, D.J., Diefes, K., Phillips, H.S., Pujara, K., Berman, B.P. et al. (2010) Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell* **17**, 510–522, <https://doi.org/10.1016/j.ccr.2010.03.017>
- 45 Toyota, M., Ahuja, N., Ohe-Toyota, M., Herman, J.G., Baylin, S.B. and Issa, J.P. (1999) CpG island methylator phenotype in colorectal cancer. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8681–8686, <https://doi.org/10.1073/pnas.96.15.8681>
- 46 Wajed, S.A., Laird, P.W. and DeMeester, T.R. (2001) DNA methylation: an alternative pathway to cancer. *Ann. Surg.* **234**, 10–20, <https://doi.org/10.1097/0000658-200107000-00003>
- 47 Ohm, J.E., McGarvey, K.M., Yu, X., Cheng, L.Z., Schuebel, K.E., Cope, L. 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, <https://doi.org/10.1038/ng1972>
- 48 Schlesinger, Y., Straussman, R., Keshet, I., Farkash, S., Hecht, M., Zimmerman, J. 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, <https://doi.org/10.1038/ng1950>
- 49 Weisenberger, D.J., Siegmund, K.D., Campan, M., Young, J., Long, T.I., Faasse, M.A. et al. (2006) CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat. Genet.* **38**, 787–793, <https://doi.org/10.1038/ng1834>
- 50 Gal-Yam, E.N., Egger, G., Iniguez, L., Holster, H., Einarsson, S., Zhang, X. et al. (2008) Frequent switching of Polycomb repressive marks and DNA hypermethylation in the PC3 prostate cancer cell line. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 12979–12984, <https://doi.org/10.1073/pnas.0806437105>
- 51 Hansen, K.D., Timp, W., Bravo, H.C., Sabuncian, S., Langmead, B., McDonald, O.G. et al. (2011) Increased methylation variation in epigenetic domains across cancer types. *Nat. Genet.* **43**, 768–775, <https://doi.org/10.1038/ng.865>
- 52 Berman, B.P., Weisenberger, D.J., Aman, J.F., Hinoue, T., Ramjan, Z., Liu, Y. et al. (2012) Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nat. Genet.* **44**, 40–46, <https://doi.org/10.1038/ng.969>
- 53 Bert, S.A., Robinson, M.D., Strbenac, D., Statham, A.L., Song, J.Z., Hulf, T. et al. (2013) Regional activation of the cancer genome by long-range epigenetic remodeling. *Cancer Cell* **23**, 9–22, <https://doi.org/10.1016/j.ccr.2012.11.006>
- 54 Feinberg, A.P. and Vogelstein, B. (1983) Hypomethylation Distinguishes Genes of Some Human Cancers from Their Normal Counterparts. *Nature* **301**, 89–92, <https://doi.org/10.1038/301089a0>
- 55 Akhtar-Zaidi, B., Cowper-Sal-lari, R., Corradin, O., Saiakhova, A., Bartels, C.F., Balasubramanian, D. et al. (2012) Epigenomic enhancer profiling defines a signature of colon cancer. *Science* **336**, 736–739, <https://doi.org/10.1126/science.1217277>
- 56 Taberlay, P.C., Achinger-Kawecka, J., Lun, A.T., Buske, F.A., Sabir, K., Gould, C.M. et al. (2016) Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. *Genome Res.* **26**, 719–731, <https://doi.org/10.1101/gr.201517.115>
- 57 Ehrlich, M. and Lacey, M. (2013) DNA hypomethylation and hemimethylation in cancer. *Adv. Exp. Med. Biol.* **754**, 31–56, [https://doi.org/10.1007/978-1-4419-9967-2\\_2](https://doi.org/10.1007/978-1-4419-9967-2_2)
- 58 Narayan, A., Ji, W., Zhang, X.Y., Marrogi, A., Graff, J.R., Baylin, S.B. et al. (1998) Hypomethylation of pericentromeric DNA in breast adenocarcinomas. *Int. J. Cancer* **77**, 833–838, [https://doi.org/10.1002\(SICI\)1097-0215\(19980911\)77:6%3c833::AID-IJC6%3e3.0.CO;2-V](https://doi.org/10.1002(SICI)1097-0215(19980911)77:6%3c833::AID-IJC6%3e3.0.CO;2-V)
- 59 Zhang, Y., Yang, L., Kucherlapati, M., Hadjipanayis, A., Pantazi, A., Bristow, C.A. et al. (2019) Global impact of somatic structural variation on the DNA methylome of human cancers. *Genome Biol.* **20**, 209, <https://doi.org/10.1186/s13059-019-1818-9>
- 60 Laird, P.W. and Jaenisch, R. (1994) DNA methylation and cancer. *Hum. Mol. Genet.* **3**, 1487–1495, <https://doi.org/10.1093/hmg/3.suppl.1.1487>
- 61 Sharma, S., De Carvalho, D.D., Jeong, S., Jones, P.A. and Liang, G. (2011) Nucleosomes containing methylated DNA stabilize DNA methyltransferases 3A/3B and ensure faithful epigenetic inheritance. *PLoS Genet.* **7**, e1001286, <https://doi.org/10.1371/journal.pgen.1001286>
- 62 You, J.S., Kelly, T.K., De Carvalho, D.D., Taberlay, P.C., Liang, G. and Jones, P.A. (2011) OCT4 establishes and maintains nucleosome-depleted regions that provide additional layers of epigenetic regulation of its target genes. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 14497–14502, <https://doi.org/10.1073/pnas.1111309108>
- 63 Widschwendter, M., Fiegl, H., Egle, D., Mueller-Holzner, E., Spizzo, G., Marth, C. et al. (2007) Epigenetic stem cell signature in cancer. *Nat. Genet.* **39**, 157–158, <https://doi.org/10.1038/ng1941>
- 64 De Carvalho, D.D., Sharma, S., You, J.S., Su, S.F., Taberlay, P.C., Kelly, T.K. et al. (2012) DNA methylation screening identifies driver epigenetic events of cancer cell survival. *Cancer Cell* **21**, 655–667, <https://doi.org/10.1016/j.ccr.2012.03.045>
- 65 Song, C.X., Szulwach, K.E., Dai, Q., Fu, Y., Mao, S.Q., Lin, L. et al. (2013) Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell* **153**, 678–691, <https://doi.org/10.1016/j.cell.2013.04.001>
- 66 Shen, L., Wu, H., Diep, D., Yamaguchi, S., D'Alessio, A.C., Fung, H.L. et al. (2013) Genome-wide Analysis Reveals TET- and TDG-Dependent 5-Methylcytosine Oxidation Dynamics. *Cell* **153**, 692–706, <https://doi.org/10.1016/j.cell.2013.04.002>
- 67 Neri, F., Incarnato, D., Krepelova, A., Rapelli, S., Pagnani, A., Zecchina, R. et al. (2013) Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. *Genome Biol.* **14**, R91, <https://doi.org/10.1186/gb-2013-14-8-r91>
- 68 Jin, C.L., Lu, Y., Jelinek, J., Liang, S.D., Estecio, M.R.H., Barton, M.C. et al. (2014) TET1 is a maintenance DNA demethylase that prevents methylation spreading in differentiated cells. *Nucleic Acids Res.* **42**, 6956–6971, <https://doi.org/10.1093/nar/gku372>



- 69 Li, X., Liu, Y., Salz, T., Hansen, K.D. and Feinberg, A. (2016) Whole-genome analysis of the methylome and hydroxymethylome in normal and malignant lung and liver. *Genome Res.* **26**, 1730–1741, <https://doi.org/10.1101/gr.211854.116>
- 70 Skvortsova, K., Masle-Farquhar, E., Luu, P.L., Song, J.Z., Qu, W., Zotenko, E. et al. (2019) DNA Hypermethylation Encroachment at CpG Island Borders in Cancer Is Predisposed by H3K4 Monomethylation Patterns. *Cancer Cell* **35**, 297e298–314e298, <https://doi.org/10.1016/j.ccell.2019.01.004>
- 71 Jeong, M., Sun, D.Q., Luo, M., Huang, Y., Challen, G.A., Rodriguez, B. et al. (2014) Large conserved domains of low DNA methylation maintained by Dnmt3a. *Nat. Genet.* **46**, 17–23, <https://doi.org/10.1038/ng.2836>
- 72 Hahn, M.A., Qiu, R., Wu, X., Li, A.X., Zhang, H., Wang, J. et al. (2013) Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. *Cell Rep.* **3**, 291–300, <https://doi.org/10.1016/j.celrep.2013.01.011>
- 73 Colquitt, B.M., Allen, W.E., Barnea, G. and Lomvardas, S. (2013) Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 14682–14687, <https://doi.org/10.1073/pnas.1302759110>
- 74 Song, C.X., Szulwach, K.E., Fu, Y., Dai, Q., Yi, C., Li, X. et al. (2011) Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat. Biotechnol.* **29**, 68–72, <https://doi.org/10.1038/nbt.1732>
- 75 Mellen, M., Ayata, P., Dewell, S., Kriaucionis, S. and Heintz, N. (2012) MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* **151**, 1417–1430, <https://doi.org/10.1016/j.cell.2012.11.022>
- 76 Wen, L., Li, X., Yan, L., Tan, Y., Li, R., Zhao, Y. et al. (2014) Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. *Genome Biol.* **15**, R49, <https://doi.org/10.1186/gb-2014-15-3-r49>
- 77 Jin, S.G., Wu, X.W., Li, A.X. and Pfeifer, G.P. (2011) Genomic mapping of 5-hydroxymethylcytosine in the human brain. *Nucleic Acids Res.* **39**, 5015–5024, <https://doi.org/10.1093/nar/gkr120>
- 78 Zhou, T., Xiong, J., Wang, M.Z., Yang, N., Wong, J.M., Zhu, B. et al. (2014) Structural Basis for Hydroxymethylcytosine Recognition by the SRA Domain of UHRF2. *Mol. Cell* **54**, 879–886, <https://doi.org/10.1016/j.molcel.2014.04.003>
- 79 Jin, S.G., Kadam, S. and Pfeifer, G.P. (2010) Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res.* **38**, e125, <https://doi.org/10.1093/nar/gkq223>
- 80 Maunakea, A.K., Chepelev, I., Cui, K. and Zhao, K. (2013) Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res.* **23**, 1256–1269, <https://doi.org/10.1038/cr.2013.110>
- 81 Khare, T., Pai, S., Koncevicius, K., Pal, M., Kriukiene, E., Liutkeviciute, Z. et al. (2012) 5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary. *Nat. Struct. Mol. Biol.* **19**, 1037–1043, <https://doi.org/10.1038/nsmb.2372>
- 82 Huang, Y., Chavez, L., Chang, X., Wang, X., Pastor, W.A., Kang, J. et al. (2014) Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1361–1366, <https://doi.org/10.1073/pnas.1322921111>
- 83 Schwartz, S., Meshorer, E. and Ast, G (2009) Chromatin organization marks exon-intron structure. *Nat. Struct. Mol. Biol.* **16**, 990–995, <https://doi.org/10.1038/nsmb.1659>
- 84 Andersson, R., Enroth, S., Rada-Iglesias, A., Wadelius, C. and Komorowski, J. (2009) Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Res.* **19**, 1732–1741, <https://doi.org/10.1101/gr.092353.109>
- 85 Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M. and Misteli, T. (2010) Regulation of Alternative Splicing by Histone Modifications. *Science* **327**, 996–1000, <https://doi.org/10.1126/science.1184208>
- 86 Verma, N., Pan, H., Dore, L.C., Shukla, A., Li, Q.V., Pelham-Webb, B. et al. (2018) TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. *Nat. Genet.* **50**, 83–95, <https://doi.org/10.1038/s41588-017-0002-y>
- 87 Manzo, M., Wirz, J., Ambrosi, C., Villasenor, R., Roschitzki, B. and Baubec, T. (2017) Isoform-specific localization of DNMT3A regulates DNA methylation fidelity at bivalent CpG islands. *EMBO J.* **36**, 3421–3434, <https://doi.org/10.15252/embj.201797038>
- 88 Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S. and Jacobsen, S.E. (2011) 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol.* **12**, R54, <https://doi.org/10.1186/gb-2011-12-6-r54>
- 89 Yu, M., Hon, G.C., Szulwach, K.E., Song, C.X., Zhang, L., Kim, A. et al. (2012) Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell* **149**, 1368–1380, <https://doi.org/10.1016/j.cell.2012.04.027>
- 90 Feeney, L., Carvalhal, V., Yu, X.C., Chan, B., Michels, D.A., Wang, Y.J. et al. (2013) Eliminating tyrosine sequence variants in CHO cell lines producing recombinant monoclonal antibodies. *Biotechnol. Bioeng.* **110**, 1087–1097, <https://doi.org/10.1002/bit.24759>
- 91 Li, J., Wu, X., Zhou, Y., Lee, M., Guo, L., Han, W. et al. (2018) Decoding the dynamic DNA methylation and hydroxymethylation landscapes in endodermal lineage intermediates during pancreatic differentiation of hESC. *Nucleic. Acids. Res.* **46**, 2883–2900, <https://doi.org/10.1093/nar/gky063>
- 92 Hon, G.C., Song, C.X., Du, T., Jin, F., Selvaraj, S., Lee, A.Y. et al. (2014) 5mC Oxidation by Tet2 Modulates Enhancer Activity and Timing of Transcriptome Reprogramming during Differentiation. *Mol. Cell*, <https://doi.org/10.1016/j.molcel.2014.08.026>
- 93 Charlet, J., Duymich, C.E., Lay, F.D., Mundbjerg, K., Dalsgaard Sorensen, K., Liang, G. et al. (2016) Bivalent Regions of Cytosine Methylation and H3K27 Acetylation Suggest an Active Role for DNA Methylation at Enhancers. *Mol. Cell* **62**, 422–431, <https://doi.org/10.1016/j.molcel.2016.03.033>
- 94 Lian, C.G., Xu, Y.F., Ceol, C., Wu, F.Z., Larson, A., Dresser, K. et al. (2012) Loss of 5-Hydroxymethylcytosine Is an Epigenetic Hallmark of Melanoma. *Cell* **150**, 1135–1146, <https://doi.org/10.1016/j.cell.2012.07.033>
- 95 Haffner, M.C., Chaux, A., Meeker, A.K., Esopi, D.M., Gerber, J., Pellakuru, L.G. et al. (2011) Global 5-hydroxymethylcytosine content is significantly reduced in tissue stem/progenitor cell compartments and in human cancers. *Oncotarget* **2**, 627–637, <https://doi.org/10.18632/oncotarget.316>
- 96 Yang, H., Liu, Y., Bai, F., Zhang, J.Y., Ma, S.H., Liu, J. et al. (2013) Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation. *Oncogene* **32**, 663–669, <https://doi.org/10.1038/nc.2012.67>
- 97 Jin, S.G., Jiang, Y., Qiu, R., Rauch, T.A., Wang, Y., Schackert, G. et al. (2011) 5-Hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. *Cancer Res.* **71**, 7360–7365, <https://doi.org/10.1158/0008-5472.CAN-11-2023>
- 98 Palomero, T., Couronne, L., Khiabani, H., Kim, M.Y., Ambesi-Impimbato, A., Perez-Garcia, A. et al. (2014) Recurrent mutations in epigenetic regulators, RHOA and FYN kinase in peripheral T cell lymphomas. *Nat. Genet.* **46**, 166–170, <https://doi.org/10.1038/ng.2873>

- 99 Abdel-Wahab, O., Mullally, A., Hedvat, C., Garcia-Manero, G., Patel, J., Wadleigh, M. et al. (2009) Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood* **114**, 144–147, <https://doi.org/10.1182/blood-2009-03-210039>
- 100 Ko, M., Huang, Y., Jankowska, A.M., Pape, U.J., Tahliliani, M., Bandukwala, H.S. et al. (2010) Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468**, 839–843, <https://doi.org/10.1038/nature09586>
- 101 Konstantin, N., Bultmann, S., Szewagierczak, A., Dufour, A., Ksienzyk, B., Schneider, F. et al. (2011) Genomic 5-hydroxymethylcytosine levels correlate with TET2 mutations and a distinct global gene expression pattern in secondary acute myeloid leukemia. *Leukemia* **25**, 1649–1652, <https://doi.org/10.1038/leu.2011.134>
- 102 Pronier, E., Almire, C., Mokrani, H., Vasanthakumar, A., Simon, A., da Costa Reis Monte Mor, B. et al. (2011) Inhibition of TET2-mediated conversion of 5-methylcytosine to 5-hydroxymethylcytosine disturbs erythroid and granulomonocytic differentiation of human hematopoietic progenitors. *Blood* **118**, 2551–2555, <https://doi.org/10.1182/blood-2010-12-324707>
- 103 Rasmussen, K.D., Jia, G., Johansen, J.V., Pedersen, M.T., Rapin, N., Bagger, F.O. 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, <https://doi.org/10.1101/gad.260174.115>
- 104 Yang, H., Ye, D., Guan, K.L. and Xiong, Y. (2012) IDH1 and IDH2 mutations in tumorigenesis: mechanistic insights and clinical perspectives. *Clin. Cancer Res.: An Off. J. Am. Assoc. Cancer Res.* **18**, 5562–5571, <https://doi.org/10.1158/1078-0432.CCR-12-1773>
- 105 Rasmussen, K.D. and Helin, K. (2016) Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev.* **30**, 733–750, <https://doi.org/10.1101/gad.276568.115>
- 106 Figueroa, M.E., Abdel-Wahab, O., Lu, C., Ward, P.S., Patel, J., Shih, A. 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, <https://doi.org/10.1016/j.ccr.2010.11.015>
- 107 Dang, L., White, D.W., Gross, S., Bennett, B.D., Bittiger, M.A., Driggers, E.M. et al. (2009) Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739–744, <https://doi.org/10.1038/nature08617>
- 108 Xu, W., Yang, H., Liu, Y., Yang, Y., Wang, P., Kim, S.H. et al. (2011) Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer Cell* **19**, 17–30, <https://doi.org/10.1016/j.ccr.2010.12.014>
- 109 Yan, H., Parsons, D.W., Jin, G., McLendon, R., Rasheed, B.A., Yuan, W. et al. (2009) IDH1 and IDH2 mutations in gliomas. *N. Engl. J. Med.* **360**, 765–773, <https://doi.org/10.1056/NEJMoa0808710>
- 110 Ward, P.S., Patel, J., Wise, D.R., Abdel-Wahab, O., Bennett, B.D., Collier, H.A. et al. (2010) The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. *Cancer Cell* **17**, 225–234, <https://doi.org/10.1016/j.ccr.2010.01.020>
- 111 Hsu, C.H., Peng, K.L., Kang, M.L., Chen, Y.R., Yang, Y.C., Tsai, C.H. et al. (2012) TET1 suppresses cancer invasion by activating the tissue inhibitors of metalloproteinases. *Cell Rep.* **2**, 568–579, <https://doi.org/10.1016/j.celrep.2012.08.030>
- 112 Thienpont, B., Steinbacher, J., Zhao, H., D'Anna, F., Kuchnio, A., Ploumakis, A. et al. (2016) Tumour hypoxia causes DNA hypermethylation by reducing TET activity. *Nature* **537**, 63–68, <https://doi.org/10.1038/nature19081>
- 113 Blaschke, K., Ebata, K.T., Karimi, M.M., Zepeda-Martinez, J.A., Goyal, P., Mahapatra, S. et al. (2013) Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **500**, 222–226, <https://doi.org/10.1038/nature12362>
- 114 Chen, J., Guo, L., Zhang, L., Wu, H., Yang, J., Liu, H. et al. (2013) Vitamin C modulates TET1 function during somatic cell reprogramming. *Nat. Genet.* **45**, 1504–1509, <https://doi.org/10.1038/ng.2807>
- 115 Peng, D., Ge, G., Gong, Y., Zhan, Y., He, S., Guan, B. et al. (2018) Vitamin C increases 5-hydroxymethylcytosine level and inhibits the growth of bladder cancer. *Clin. Epigenetics* **10**, 94, <https://doi.org/10.1186/s13148-018-0527-7>
- 116 Gustafson, C.B., Yang, C., Dickson, K.M., Shao, H., Van Booven, D., Harbour, J.W. et al. (2015) Epigenetic reprogramming of melanoma cells by vitamin C treatment. *Clin. Epigenetics* **7**, 51, <https://doi.org/10.1186/s13148-015-0087-z>
- 117 Kuo, K.C., McCune, R.A., Gehrke, C.W., Midgett, R. and Ehrlich, M. (1980) Quantitative reversed-phase high performance liquid chromatographic determination of major and modified deoxyribonucleosides in DNA. *Nucleic. Acids. Res.* **8**, 4763–4776, <https://doi.org/10.1093/nar/8.20.4763>
- 118 Vryer, R. and Saffery, R. (2017) What's in a name? Context-dependent significance of 'global' methylation measures in human health and disease. *Clin. Epigenetics* **9**, 2, <https://doi.org/10.1186/s13148-017-0311-0>
- 119 Stirzaker, C., Taberlay, P.C., Statham, A.L. and Clark, S.J. (2014) Mining cancer methylomes: prospects and challenges. *Trends in Genetics: TIG* **30**, 75–84, <https://doi.org/10.1016/j.tig.2013.11.004>
- 120 Fraga, M.F., Rodriguez, R. and Canal, M.J. (2000) Rapid quantification of DNA methylation by high performance capillary electrophoresis. *Electrophoresis* **21**, 2990–2994, [https://doi.org/10.1002/1522-2683\(20000801\)21:14%3c2990::AID-ELPS2990%3e3.0.CO;2-I](https://doi.org/10.1002/1522-2683(20000801)21:14%3c2990::AID-ELPS2990%3e3.0.CO;2-I)
- 121 Friso, S., Choi, S.W., Dolnikowski, G.G. and Selhub, J. (2002) A method to assess genomic DNA methylation using high-performance liquid chromatography/electrospray ionization mass spectrometry. *Anal. Chem.* **74**, 4526–4531, <https://doi.org/10.1021/ac020050h>
- 122 Karimi, M., Johansson, S., Stach, D., Corcoran, M., Grandier, D., Schalling, M. et al. (2006) LUMA (Luminometric Methylation Assay)—a high throughput method to the analysis of genomic DNA methylation. *Exp. Cell Res.* **312**, 1989–1995, <https://doi.org/10.1016/j.yexcr.2006.03.006>
- 123 Lisanti, S., Omar, W.A., Tomaszewski, B., De Prins, S., Jacobs, G., Koppen, G. et al. (2013) Comparison of methods for quantification of global DNA methylation in human cells and tissues. *PLoS One* **8**, e79044, <https://doi.org/10.1371/journal.pone.0079044>
- 124 Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H. et al. (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* **133**, 523–536, <https://doi.org/10.1016/j.cell.2008.03.029>
- 125 Clark, S.J., Harrison, J., Paul, C.L. and Frommer, M. (1994) High sensitivity mapping of methylated cytosines. *Nucleic. Acids. Res.* **22**, 2990–2997, <https://doi.org/10.1093/nar/22.15.2990>
- 126 Kelly, T.K., Liu, Y., Lay, F.D., Liang, G., Berman, B.P. and Jones, P.A. (2012) Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome Res.* **22**, 2497–2506, <https://doi.org/10.1101/gr.143008.112>

- 127 Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L. et al. (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* **37**, 853–862, <https://doi.org/10.1038/ng1598>
- 128 Serre, D., Lee, B.H. and Ting, A.H. (2010) MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. *Nucleic. Acids. Res.* **38**, 391–399, <https://doi.org/10.1093/nar/gkp992>
- 129 Pidsley, R., Zotenko, E., Peters, T.J., Lawrence, M.G., Risbridger, G.P., Molloy, P. et al. (2016) Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol.* **17**, 208, <https://doi.org/10.1186/s13059-016-1066-1>
- 130 Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V.K., Attwood, J., Burger, M. et al. (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat. Genet.* **38**, 1378–1385, <https://doi.org/10.1038/ng1909>
- 131 Tian, Y., Morris, T.J., Webster, A.P., Yang, Z., Beck, S., Feber, A. et al. (2017) ChAMP: updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics* **33**, 3982–3984, <https://doi.org/10.1093/bioinformatics/btx513>
- 132 Bock, C., Tomazou, E.M., Brinkman, A.B., Muller, F., Simmer, F., Gu, H. et al. (2010) Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nat. Biotechnol.* **28**, 1106–1114, <https://doi.org/10.1038/nbt.1681>
- 133 Smallwood, S.A., Lee, H.J., Angermueller, C., Krueger, F., Saadeh, H., Peat, J. et al. (2014) Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat. Methods* **11**, 817–820, <https://doi.org/10.1038/nmeth.3035>
- 134 Angermueller, C., Clark, S.J., Lee, H.J., Macaulay, I.C., Teng, M.J., Hu, T.X. et al. (2016) Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat. Methods* **13**, 229–232, <https://doi.org/10.1038/nmeth.3728>
- 135 Pott, S. (2017) Simultaneous measurement of chromatin accessibility, DNA methylation, and nucleosome phasing in single cells. *Elife* **6**, e23203, <https://doi.org/10.7554/eLife.23203>
- 136 Wang, Y., Wang, A., Liu, Z., Thurman, A.L., Powers, L.S., Zou, M. et al. (2019) Single-molecule long-read sequencing reveals the chromatin basis of gene expression. *Genome Res.* **29**, 1329–1342, <https://doi.org/10.1101/gr.251116.119>
- 137 Grunau, C., Clark, S.J. and Rosenthal, A. (2001) Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic. Acids. Res.* **29**, E65–65, <https://doi.org/10.1093/nar/29.13.e65>
- 138 Tanaka, K. and Okamoto, A. (2007) Degradation of DNA by bisulfite treatment. *Bioorg. Med. Chem. Lett.* **17**, 1912–1915, <https://doi.org/10.1016/j.bmcl.2007.01.040>
- 139 Schutsky, E.K., DeNizio, J.E., Hu, P., Liu, M.Y., Nabel, C.S., Fabyanic, E.B. et al. (2018) Nondestructive, base-resolution sequencing of 5-hydroxymethylcytosine using a DNA deaminase. *Nat. Biotechnol.*, <https://doi.org/10.1038/nbt.4204>
- 140 Zeng, H., He, B., Xia, B., Bai, D., Lu, X., Cai, J. et al. (2018) Bisulfite-Free, Nanoscale Analysis of 5-Hydroxymethylcytosine at Single Base Resolution. *J. Am. Chem. Soc.* **140**, 13190–13194, <https://doi.org/10.1021/jacs.8b08297>
- 141 Raiber, E.A., Beraldi, D., Ficiz, G., Burgess, H.E., Branco, M.R., Murat, P. et al. (2012) Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. *Genome Biol.* **13**, R69, <https://doi.org/10.1186/gb-2012-13-8-r69>
- 142 Xia, B., Han, D., Lu, X., Sun, Z., Zhou, A., Yin, Q. et al. (2015) Bisulfite-free, base-resolution analysis of 5-formylcytosine at the genome scale. *Nat. Methods* **12**, 1047–1050, <https://doi.org/10.1038/nmeth.3569>
- 143 Zhu, C., Gao, Y., Guo, H., Xia, B., Song, J., Wu, X. et al. (2017) Single-Cell 5-Formylcytosine Landscapes of Mammalian Early Embryos and ESCs at Single-Base Resolution. *Cell Stem Cell* **20**, 720e725–731e725, <https://doi.org/10.1016/j.stem.2017.02.013>
- 144 Liu, Y., Siejka-Zielinska, P., Velikova, G., Bi, Y., Yuan, F., Tomkova, M. et al. (2019) Bisulfite-free direct detection of 5-methylcytosine and 5-hydroxymethylcytosine at base resolution. *Nat. Biotechnol.* **37**, 424–429, <https://doi.org/10.1038/s41587-019-0041-2>