Clinical Epigenetics

REVIEW

Gene body methylation in cancer: molecular mechanisms and clinical applications

Qi Wang, Fei Xiong, Guanhua Wu, Wenzheng Liu, Junsheng Chen, Bing Wang and Yongjun Chen[*]

Abstract

DNA methylation is an important epigenetic mechanism that regulates gene expression. To date, most DNA methylation studies have focussed on CpG islands in the gene promoter region, and the mechanism of methylation and the regulation of gene expression after methylation have been clearly elucidated. However, genome-wide methylation studies have shown that DNA methylation is widespread not only in promoters but also in gene bodies. Gene body methylation is widely involved in the expression regulation of many genes and is closely related to the occurrence and progression of malignant tumours. This review focusses on the formation of gene body methylation patterns, its regulation of transcription, and its relationship with tumours, providing clues to explore the mechanism of gene body methylation in regulating gene transcription and its significance and application in the field of oncology.

Keywords: Epigenetics, Gene body methylation, Histone methylation, Transcription regulation, Cancer

Introduction

In the 1970s, Riggs AD and Holliday R proposed that methylation of the 5′-terminal cytosine (5′-C) of CpG dinucleotides is an important epigenetic marker of gene silencing that can cause changes in chromatin structure and DNA stability without altering the DNA sequence, thereby regulating gene expression [1, 2]. By the 1980s, it was found that CpG sites are heterogeneously distributed in the genome, with CpG sites tending to cluster in higher densities at promoter regions than throughout the rest of the genome, which has led to the definition of the term CpG islands (CGIs) [3], while the remaining 40–50% is scattered in intragenic regions and 30–40% in intergenic regions [4]. In the human genome, there are approximately 28,890 CGIs, most of which are located in the promoter regions of functional genes and are mostly unmethylated [3–5]. Approximately 9000 CGIs are located in the gene body regions, of which 30–40% are methylated [5–7].

With the rise of technologies such as genomic bisulfite sequencing, genome-wide DNA methylation mapping has been possible, and more extensive DNA methylation in the gene body region has attracted attention. Gene body methylation, which refers to the methylation of CpG sites in the transcribed regions of genes, including the methylation of scattered CpG sites and CGIs [8], was first described in Arabidopsis (Brassicaceae/Kale) in 2005 [9]. In 2006, the first genome-wide DNA methylation profile of Arabidopsis was described, showing that pericentromeric heterochromatin, repetitive sequences and regions producing small interfering RNAs around Arabidopsis thaliana mitoses are hypermethylated, and more than one-third of expressed genes are hypermethylated within transcribed regions, while only approximately 5% of genes exhibit methylation within promoter regions [10]. One year later, Zilberman discovered a significant correlation between gene body methylation and gene transcription in Arabidopsis [11]. Subsequent studies found that gene body methylation in animal genomes is similar to that in plants, preferentially occurring in exons of moderately, constitutively expressed and evolutionarily

*Correspondence: Yjchen@tjh.tjmu.edu.cn

Department of Biliary-Pancreatic Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, No. 1095 Jiefang Road, Wuhan 430074, Hubei Province, China

Wang *et al. Clinical Epigenetics*   (2022) 14:154

Page 2 of 14

conserved housekeeping genes and having the same role in regulating transcription [12–15]. In human genomes, Rauch et al. and Maunakea et al. confirmed that methylated gene body regions correlated with higher levels of gene transcription [4, 5]. Besides, gene body methylation can affect processes such as histone modification [16], alternative splicing [17], and spurious transcription [18], discussed later in this review. At the same time, numerous studies have shown that abnormal methylation of gene bodies is closely related to gene expression, growth and differentiation, development, treatment and prognosis of tumours. In this review, we will discuss the distribution pattern of gene body methylation, its role and its relationship with tumours and explore the clinical application of gene body methylation and the urgent issues to be solved.

Mechanism of gene body methylation patterning generation
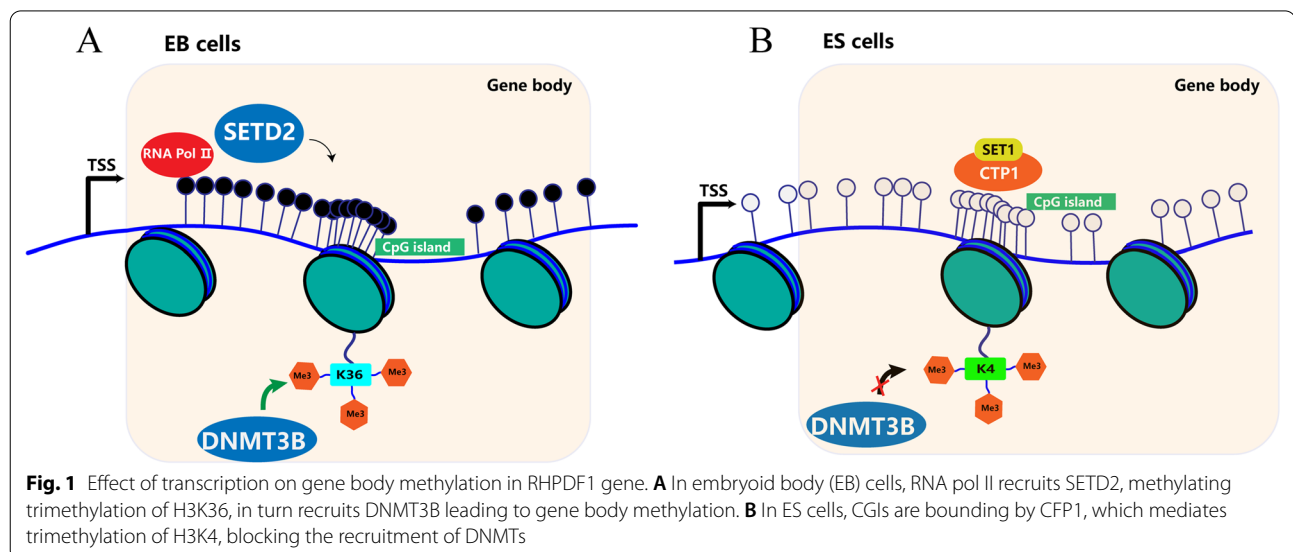
Active transcription promotes gene body methylation

Although DNA methylation is commonly thought to lead to gene silencing, in the case of genomic methylation, it appears to be regulated by active transcription while affecting gene transcription. Following genetic and acquired genomic rearrangements, aberrant methylation of CGIs occurs when an unmethylated CGI is aberrantly recombined into another transcriptionally active gene in vivo [19, 20]. Thus, methylation of some CGIs may be a consequence rather than a cause of transcription. It has been shown that in the RHBDF1 gene without a promoter, the level of methylation of the gene body was significantly lower than in the RHBDF1 gene with a promoter. Besides, H3K36me3, a histone modification, is closely associated with gene body methylation through the recruitment of DNA methyltransferase 3B (DNMT 3B), and its distribution level can indirectly reflect gene body methylation levels [21]. It has been shown that H3K36me3 in the RHBDF1 gene was not enriched at the transcription start site and gradually increased throughout the gene body, reaching the maximum level near the transcription termination site. This phenomenon suggests that the methylation of CGIs within the RHBDF1 gene is dependent on promoter-mediated transcriptional processes. Overall, it has been shown that whether a specific alternative promoter or intergenic CGI becomes methylated depends on its chromosomal context and, in particular, on whether transcription passes through the CGI at some point [16] (Fig. 1A).

However, in ES cells, the CGIs through which transcription extends are not methylated. This is because, in ES cells, CGIs are usually bound by CpG binding factors such as CFP1, a component of the Set1 complex that mediates trimethylation of H3K4, a marker known to block the recruitment of DNMT 3A/3B/3L [22–25] (Fig. 1B). As cells differentiate, the H3K4me3 mark within the gene is gradually lost, and the transcriptional process gradually begins to mediate the methylation of CGIs within the gene [16].

In addition, while transcription can drive the methylation silencing of CGIs within a gene, it is not always sufficient to drive methylation of alternative promoter CGIs, which themselves act as strong initiators of transcription. Danuta et al. showed that the relative strengths of transcription initiation from the intragenic CGI and the promoter CGI of the gene within which it lies play an important role in determining the status of the intragenic

Fig. 1 Effect of transcription on gene body methylation in RHPDF1 gene. **A** In embryoid body (EB) cells, RNA pol II recruits SETD2, methylating trimethylation of H3K36, in turn recruits DNMT3B leading to gene body methylation. **B** In ES cells, CGIs are bounding by CFP1, which mediates trimethylation of H3K4, blocking the recruitment of DNMTs

Wang *et al. Clinical Epigenetics*   (2022) 14:154

Page 3 of 14

CGI. There is speculation that CGIs that are longer [26] and have higher CpG density [27] may have stronger transcriptional activity. An alternative promoter CGI will not necessarily be silenced by elongating transcripts from a linked, but weaker, CGI promoter. These CGIs may perform additional biologically important functions.

Interaction between gene body methylation and histone methylation

The majority of the promoter CGIs (∼95%) and alternative promoter CGIs (∼85%) are unmethylated (≤45% methylation). Their hypomethylated state is associated with H3K4me3 enrichment and lack of H3K36me3. In contrast, the majority (∼65%) of gene body CGIs are DNA-methylated (≥55% methylation). The hypermethylated states of these CGIs are associated with traversing transcription elongation, H3K36me3 enrichment, and lack of H3K4me3 and H3K27me3 [16]. This phenomenon suggests a close link between gene body methylation and histone marks.

Roles of H3K4me3, H3K9me3 and H4K20me3 in DNA methylation

H3K4 is normally distributed on promoter CGIs and is regulated by the competitive binding of the CXXC domain of H3K4 methyltransferase and the ADD domain of DNMTs [22, 28]. Thus, DNA methylation and histone H3K4 methylation are negatively correlated. In the gene body, the vast majority of CGIs are hypermethylated, and thus, mostly lack H3K4me3 marks.

H3K9me3 binds primarily to heterochromatin and mediates DNA methylation of heterochromatin regions by the methyltransferase DNMT3A/B: 1. DNMT3A/B is recruited to heterochromatin by interacting with heterochromatin protein 1 (HP1) bound to H3K9me3 [29]; 2. DNMT3A/B also interacts directly with H3K9 methyltransferases. DNMT3A/B have been shown to interact with Suv39h1 [30] and Setdb1 [31] via their ADD domains. Furthermore, DNMT3A interacts with the euchromatin-associated H3K9 methyltransferase G9a/GLP via the chromatin domain protein MPP8 [32]. Additionally, a direct interaction between the C-terminal catalytic domains of DNMT3A/B and the ankyrin repeat domains of G9a, which are themselves able to bind to H3K9 methylation, has been shown to play roles in de novo methylation [33]. However, H3K9me is not clearly associated with gene body methylation, which was first found in plants [34]. Genome-wide methylation sequencing revealed that in suvh4, suvh5 and suvh6 mutant plants, gene body methylation levels were almost unaffected [35]. Therefore, the relationship between gene body methylation and H3K9me has been rarely studied.

H4K20me3 is a repressive mark found in constitutive heterochromatin and at imprinted genes where it is selectively enriched on the DNA-methylated allele [36–38]. In promoter regions, H4K20me3 represses gene expression as does DNA methylation [39], but the interaction between the two is not yet known. In heterochromatin regions, DNMT1 directly recognizes H4K20me3 through the first bromo-adjacent-homology domain, which in turn mediates the chromatin association and enzymatic activation of DNMT1 [40]. In the gene body regions, however, H3K20me3 is antagonistic to DNA methylation. The genetic regulator plant homologous structural domain finger protein 6 (Phf6) binds to the demethylated rDNA gene body and recruits the histone methyltransferase suv4-20h2 to establish the histone modification H4K20me3, which ultimately represses rDNA transcription. Gene body methylation prevents the recruitment of PHF6 and thus ensures rDNA transcription [41]. However, a direct interaction between H4K20me3 and gene body methylation has not been reported.

H3K36me3 and gene body methylation

H3K36me3 is mainly marked on the gene body and is formed by the histone methylation enzyme SETD2 and in conjunction with RNA Pol II catalyses the methylation of H3K36me0 [42], which is closely related to DNA de novo methylation. The de novo methylation of DNA is mainly initiated by the DNMT3 family [43], of which DNMT3B1 mainly binds to the gene body [21]. DNMT3 contains a PWWP domain that specifically recognizes H3K36me3 and is recruited to a highly abundant H3K36me3-labelled genomic region, causing the CpG site to undergo de novo methylation. Baubec et al. [21] knocked down the PWWP domain of DNMT3 or downregulated SETD2 in mouse embryonic stem (ES) cells and found that the accumulation of H3K36me3 was blocked and the level of gene body methylation decreased. This finding indicates that H3K36me3 plays an important role in recruiting DNMT3 and promoting gene body DNA methylation. H3K36me2 primarily recruits DNMT3A to catalyse and maintain DNA methylation in the intergenic region. Deletion of H3K36me2 results in redistribution of DNMT3A to H3K36me3-tagged gene bodies and reduces intergenic DNA methylation levels, suggesting that the PWWP structural domain of DNMT3A has dual recognition of H3K36me2/3 with a higher binding affinity towards H3K36me2 [44] and suggesting that posttranscriptional methylation of histone H3K36 plays a crucial role in the regulation of methylation levels in gene bodies.

H3K27me3 and gene body methylation

H3K27me3 is an epigenetic marker that inhibits the transcriptional elongation of RNA Pol II and promotes

Wang *et al. Clinical Epigenetics*   (2022) 14:154

Page 4 of 14

chromatin densification. H3K27me3 whole gene mapping shows that H3K27me3 is widely enriched in the gene body [45]. Removal of H3K27me3 from the gene body by JMJD3 demethylase promotes RNA Pol II elongation and activates the SETD2 complex, allowing H3K36me3 to accumulate in the gene body [46]. The relationship between H3K27 methylation and gene body methylation remains unclear, although H3K27me3 was shown to be antagonistic to H3K36me3. It has been found that H3K27me3 and DNA methylation occurred simultaneously in most regions of low CpG density in the gene body and only in CGIs were the two marks mutually exclusive. However, in DNMT knockout TKO cells, in addition to the high-density CpG regions where H3K27me3 markers were increased, large areas of H3K27me3 enrichment appeared in most regions of the genome [45]. This suggests that the two markers have an antagonistic effect on a larger scale of the genome in addition to a local antagonistic effect in the high-density CpG region. However, it has also been shown that gene body methylation coexists with H3K27me3 to maintain the chromatin dense state and repress transcription [47]. In conclusion, the relationship and interaction mechanism between the two have not been elucidated, and further experimental studies are still needed.

In conclusion, there is an inextricable relationship between histone modifications and DNA methylation, with histone modifications regulating DNA methylation and DNA methylation influencing histone modifications and acting further through them.

Functions of gene body methylation

DNA methylation near the transcription start site usually has a repressive effect on gene expression, but methylation of the gene body has the opposite effect [48–52]. This phenomenon is known as the "DNA methylation paradox" [53], and recent studies have revealed a more complex relationship between genomic methylation and transcription.

Preventing spurious transcription initiation

Spurious transcription is the process by which RNA polymerase binds to an ectopic promoter rather than a normal promoter to produce an abnormal transcript [54]. Some of the abnormal transcripts are polyadenylated and stabilized to produce abnormal proteins, which in turn affect molecular mechanisms, including gene expression regulation [55], miRNA targeting [56] and truncated protein generation [57, 58], leading to tumour cell heterogeneity and tumour susceptibility [18]. It has been found that Dnmt3b anchoring and binding to H3K36me3 in the gene body region inhibits RNA Pol II from binding to cryptic transcription initiation downstream of normal promoters by methylating the bases at abnormal transcription start sites [18]. Lorincz et al. [59] found that the density of RNA Pol II in the methylated region of the gene body is significantly reduced. Rapelli et al. [18] also found that in Dnmt3b−/− ES cells, a significant amount of RNA is transcribed from exons of the gene body. These findings confirm, both positively and negatively, that gene body methylation plays an important role in maintaining transcriptional fidelity, which prevents abnormal intragenic transcription by blocking the binding of RNA Pol II to ectopic transcription initiation within the gene body (Fig. 2).
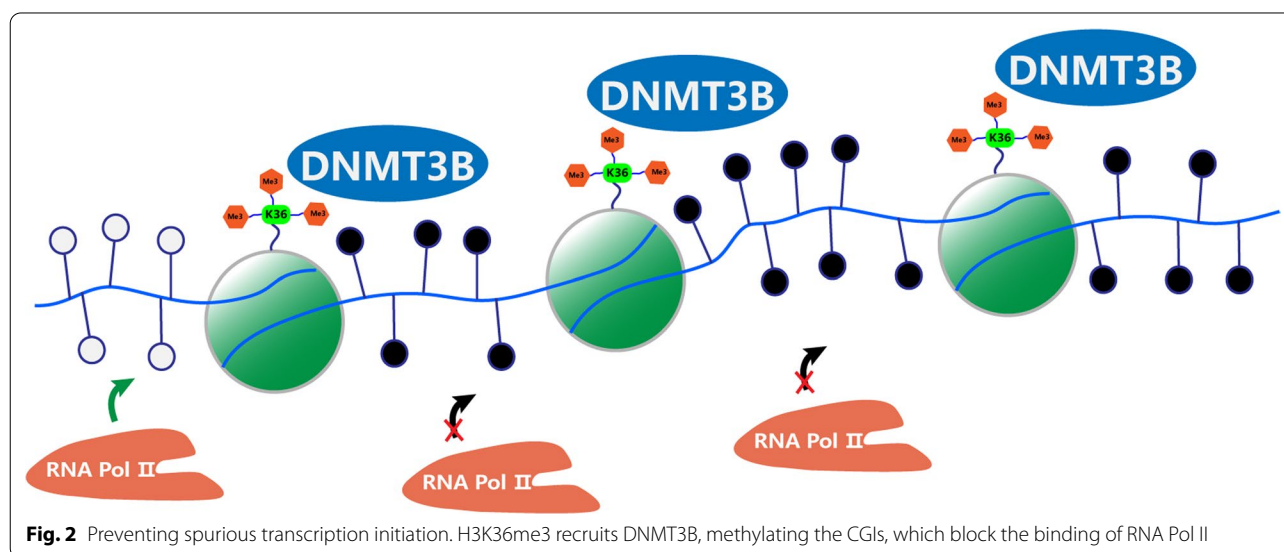

Fig. 2 Preventing spurious transcription initiation. H3K36me3 recruits DNMT3B, methylating the CGIs, which block the binding of RNA Pol II

Wang *et al. Clinical Epigenetics*   (2022) 14:154

Page 5 of 14

Inhibition of transcriptional elongation

Acetylation of the N-terminal lysine of histones is closely associated with transcriptional activation, and this modification neutralizes the positive charge of the lysine, thereby inducing a change in nucleosome conformation and making it easier for nucleosomes to participate in transcription [60, 61]. Gene body methylation induces hypoacetylation of histone lysine, promotes the formation of dense chromatin structures, prevents binding of TFs, and reduces the efficiency of RNA Pol II elongation [59].

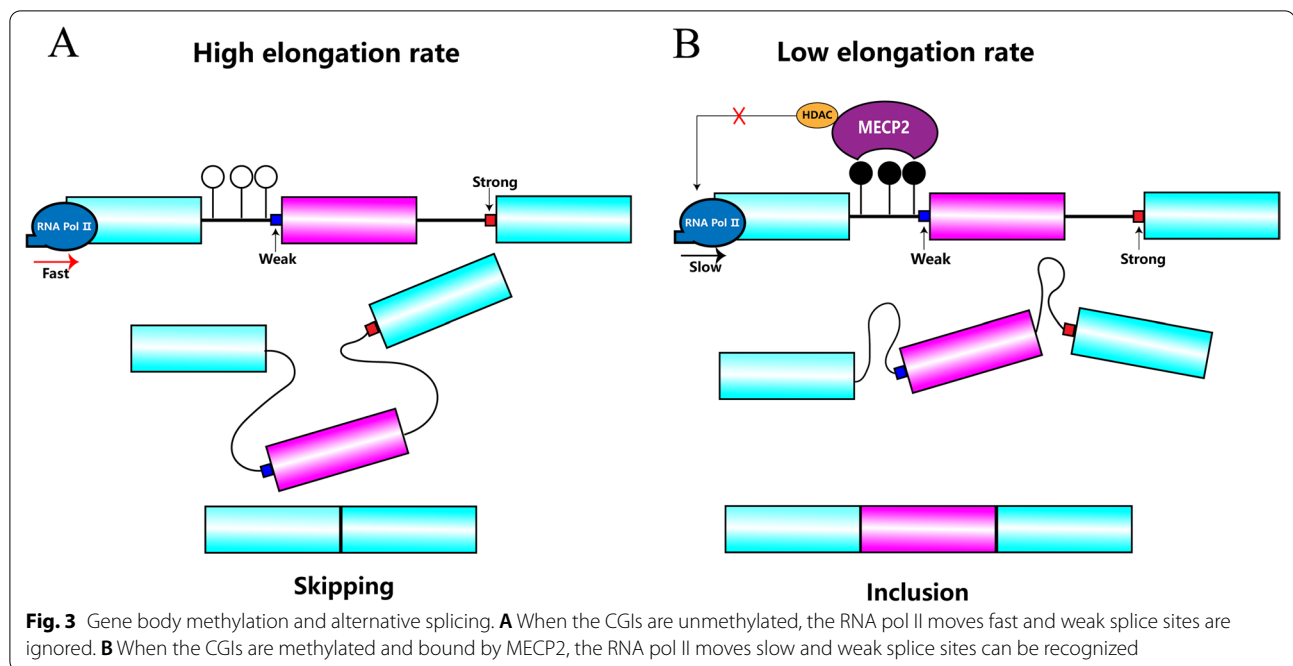### Regulating alternative splicing

In eukaryotes, pre-messenger RNAs (pre-mRNAs) undergo splicing, a process of maturation by which large intervening sequences (introns) are removed, leaving a mature transcript composed only of exons spliced together [17]. This process serves as a crucial regulatory step of gene expression and transcriptome diversification, as almost all genes produce alternative transcripts with differing exon compositions [62]. A higher level of methylation was found in the region near the exon start site and the termination site, suggesting that DNA methylation is involved in the process of exon alternative splicing. The higher the exon methylation density (total exon methylation level divided by the length of the corresponding exon), the higher the level of inclusion and expression of the exons in the transcript [63].

Two mechanisms that are not mutually exclusive have been proposed to explain how epigenetic modifications work: the kinetic model and the recruitment model. In the kinetic model, epigenetic modifications influence the kinetics of transcriptional elongation, which in turn affects splicing. This includes affecting the pace at which splice sites and regulatory sequences emerge in nascent precursor mRNAs during transcription, the rate of movement of RNA Pol II, and the open state of chromatin [64]. During transcription, when RNA Pol II moves faster, weaker splice sites are skipped, and exons at the original weak sites are skipped (Fig. 3A). When RNA Pol II moves more slowly, there is sufficient quantity and time to recognize the weaker splice sites, thus facilitating the inclusion of the corresponding exons (Fig. 3B). For example, the DNA-binding protein CTCF, which recognizes a specific unmethylated DNA sequence, binds to DNA and reduces the transcriptional elongation rate of RNA Pol II, thus facilitating exon inclusion; when the DNA is methylated, CTCF is unable to bind to it, the transcriptional elongation rate of RNA Pol II becomes faster, and exons are skipped [65]. In addition, the DNA-binding protein MECP2 recognizes and binds to methylated DNA and then recruits histone deacetylases, which reduce the rate of RNA Pol II movement and thus facilitate exon inclusion [66]. In the recruitment model, splicing regulation is accomplished through the binding of epigenetic modification sites by adaptor proteins and the recruitment of splicing factors [67]. For example, histone H3K36me3 can recruit splicing factors PTB [68] and SRSF1 [69] to regulate alternative splicing. In addition, DNA methylation can cause histone H3K9me3 to bind to heterochromatin protein 1 (HP1), which recruits the splicing factor serine- and arginine-rich splicing factor 3 (SRSF3) to the methylated region of the gene to achieve alternative splicing; when HP1 binds to an alternative exon, it leads to exon skipping, and when HP1 binds to an intron immediately upstream of the alternative exon, it promotes exon containment [70].

Enhancer methylation

It has been found that demethylation of the genome results in the appearance of many new active enhancers, evenly distributed across gene bodies, and mediates the upregulation of gene expression, suggesting that gene body methylation silences enhancer function [71]. This is mainly associated with reduced chromatin accessibility and reduced binding of transcription factors (TF) [72]. The level of methylation at TF binding sites is negatively correlated with TF expression, and TF binding induces demethylation of binding sites, so enhancer methylation may be a consequence of reduced TF expression [72]. For example, the binding of glucocorticoid receptors to tissue-specific enhancers induces DNA demethylation, leading to enhancer activation [73]. However, the relationship between DNA methylation and enhancer activity remains incompletely understood, and it has been shown that high levels of DNA methylation coexist with H3K27ac in conventional and super-enhancers and that these regions are generally located outside of the TF binding site and therefore do not inhibit enhancer activity but rather stabilize DNA and define demethylated TF binding sites [74]. In addition, bioinformatic analyses have shown that CGIs themselves exhibit enhancer histone modifications and show greater transcription factor binding [75]. Such signatures have been identified at a gene body CGI within Kdm6b, which exhibits H3K4me1 and loops to the promoter CGI to enhance Kdm6b expression [76]. In addition, the distribution of enhancer CGIs plays an important role in enhancer function; enhancers enriched in CGIs (i) are more highly conserved, (ii) are more highly enriched in R loops and 3D genomic contacts, (iii) on average provide stronger signals in functional assays, and (iv) are prone to hypermethylation in cancer [77].

Fig. 3 Gene body methylation and alternative splicing. **A** When the CGIs are unmethylated, the RNA pol II moves fast and weak splice sites are ignored. **B** When the CGIs are methylated and bound by MECP2, the RNA pol II moves slow and weak splice sites can be recognized

Gene body methylation and cell differentiation

Recent studies have revealed that many genes specifically expressed in tissues and involved in encoding key regulators of morphogenesis and organ development contain genomic CGIs, which may function as cis-acting elements controlling tissue-specific transcription [78]. By comparing human embryonic stem cells with differentiated cells, such as human neural progenitors and fibroblasts, significant changes in gene body CGI modifications were found to correlate with differentiation-induced gene expression. In embryonic stem cells, intragenic CGIs are usually bivalently modified by H3K4me3 and H3K27me3 and are not methylated. In differentiated cell gene bodies, however, there is a lack of bivalently modified CGIs, and some CGIs are monovalently modified by high levels of H3K27me3 without methylation; approximately one-third are hypermethylated, which is associated with gene activation, and these hypermethylated genes play an important role in growth and development [78]. This is associated with gene activation, and these hypermethylated genes play important roles in growth and development. For example, the differentiation of pluripotent stem cells into neuronal precursor cells is mediated by the binding of the proneural factor NEUROD1 (a transcription factor that induces differentiation into neuronal cells) to hypermethylated regions in specific gene bodies [79]. These results suggest that during differentiation, gene body CGIs regulate gene expression by converting bivalent histone modification to DNA methylation at poised developmental genes or by gaining H3K27me3

marks to further silence genes, the expression of which is not needed. Cell-type-specifically expressed transcription factors and their target genes may take advantage of epigenetic modifications of gene body CGIs to generate a spectrum of differentiation from embryonic stem cells [78]. It is thus clear that gene body methylation is closely linked to cell development and differentiation and plays an important developmental regulatory role.

Gene body methylation also plays an important role in the aberrant differentiation of tumour cells. In prostate cancer cells, DNMT1 is reduced to promote epithelial–mesenchymal transition (EMT) and conversion to tumour stem cells by promoting PKCα expression, which in turn promotes prostate cancer growth and metastasis [80]. In triple-negative breast cancer (TNBC), DNMT1 promotes the EMT required for metastasis by regulating the methylation levels of epithelial markers [81]. Another study found that HCT116 human colorectal cancer cells inactivated with DNMT1 and DNMT3B (DKO cells) showed increased expression of mesenchymal markers and decreased expression of epithelial markers, with marked changes in cell morphology and epithelial–mesenchymal transition. The splicing factors ESRP1, RBFOX2, and NBML1, which affect the composition of CD44 exons, were also downregulated in DKO cells, suggesting that the development of EMT is associated with DNA methylation and its resultant CD44 alternative exon skipping [57]. These results suggest that DNA methyltransferase deficiency plays an important role in EMT and has a significant impact on tumour hypofractionation

Wang *et al. Clinical Epigenetics* (2022) 14:154

Page 7 of 14

and growth invasion. However, this contradicts the role of the antitumor DNA methyltransferase inhibitor drugs, and the current study does not clearly identify the region of genomic demethylation. Further studies are needed to investigate the mechanism of how demethylation induces tumour dedifferentiation.

Gene body methylation and cancer

Tumour repeat element hypomethylation

The majority of CpG sites are found in repeat elements (REs) that occupy two-thirds of the genome, including interspersed repeats and tandem repeats [82, 83]. Repeated elements play a role in chromosome structure formation [84], gene transcription [85, 86] and nuclear structure [86]. In tumours, more than 2/3 of tumour genome-wide hypomethylation occurs in REs [87] (Table 1). REs hypomethylation can lead to reduced resistance of the genome to DNA damage [88], affect the stability of the mitogen and surrounding regions [89] and promote chromosomal rearrangements [90]. This can affect genomic stability and even induce oncogene activation. Furthermore, RE hypomethylation seems to be a unique feature of tumours that is not found in precancerous lesions or peripheral blood [91, 92] and therefore has potential use in cancer diagnosis and prognosis. However, the detailed mechanism by which RE methylation affects mitophagy function remains an urgent question. In addition, RE hypomethylation is also associated with hereditary diseases, neurodevelopmental, neurodegenerative, and neuropsychiatric disorders as well as ageing [93].

Gene body methylation and tumorigenesis

Huang et al. found that intestinal epithelial saprophytic cells exhibit aberrant hypermethylation in the early stages, but global gene body hypomethylation occurs in the transition from intestinal epithelial saprophytic cells to cancer cells [107]. This suggests that gene body methylation plays an important role in the development of tumours. In normal cells, METTL7A is a key protein for lipid metabolism and the generation of functional organelles. EZH2 interacts with the CpG sites of the METTL7A gene body in thyroid cancer, recruiting MBD2 and rejecting RNA pol II, resulting in METTL7A silencing, which in turn mediates abnormal tumour metabolism [108]. Significant gene body CGI hypomethylation was observed in hepatitis B virus X (HBx)-infected livers. HBx promotes hepatocarcinogenesis by recruiting histone deacetylase 1 (HDAC1) to repress the Dnmt3L and Dnmt3a promoters, inducing gene body CGI hypomethylation and downregulation of associated developmental regulators [109]. Lai et al. found that in lymphoma cells with high expression of BCL6, CGI within the first intron of the BCL6 locus was hypermethylated. The transcription factor CTCF, which blocks the enhancer, is usually bound to unmethylated DNA. During lymphomagenesis, hypermethylation of the first intron DNA blocks CTCF binding, thereby allowing the enhancer to exert its effects and increase BCL6 expression, promoting the development of lymphoma [110]. Cheung et al. [111] found that in testicular malignancies, hypermethylation of an intron region spanning miR-199a resulted in downregulation of its expression, which in turn led to abnormal upregulation of the expression of its target protein podocalyxin-like protein 1 (PODXL), affecting the aggressive behaviour of testicular tumours. The inositol triphosphate 3-kinase A gene (ITPKA) is a potentially oncogenic gene that is upregulated in a variety of cancers. Wang et al. found that the gene body CGI of ITPKA competes with the promoter to bind the SP1 transcription factor and that methylation of the gene body promotes SP1 binding to the promoter and thus ITPKA expression. In lung cancer, the methylation of ITPKA occurs in in situ carcinoma and increases with the level of invasion [112].

Table 1 Effect of REs hypomethylation on tumours

REs	Tumours	Effect	References
LINE-1	Lung cancer	Cancer stage and prognosis	[94]
Alu	Thyroid cancer	Metastasis and prognosis	[95]
LINE and Alu Yb8	Bladder cancer	Cancer stage or disease outcome	[96]
Sat2, Satα, and Alu	Liver cancer	Gain of chromosome 1q and diagnosis of cancer	[89, 97]
Satα	Nephroblastoma	Loss of chromosome 16q Gain of chromosome 1q	[98]
Sat2, Satα, LINE1 and Alu	Ovarian epithelial tumour	Prognosis and biomarkers	[99–101]
LINE-1and Alu	Breast cancer	Risk, progression, and type of cancer	[102, 103]
LINE-1 and Alu	Gastrointestinal tumours	Cancer stage and prognosis	[104–106]

Wang *et al. Clinical Epigenetics*   (2022) 14:154

Page 8 of 14

In addition, HOX transcript antisense RNA (HOTAIR), a widely studied long-stranded noncoding RNA (lncRNA), was found to be significantly overexpressed in different cancer types with different genetic profiles. It interacts with microRNAs (miRNAs) as a competing RNA to induce dysregulation of key genes, thereby driving important cancer phenotypes. In a variety of tumours, CDK9-mediated phosphorylation of RNA Pol II Ser2 and MLL1-mediated regulation of H3K4me3 allow for hypermethylation of CGIs in exons of the HOTAIR gene, facilitating the transcriptional elongation process and promoting HOTAIR overexpression, which in turn mediates tumour development [113]. McGuire et al. found that lymphocyte-specific gene CARD11 gene body methylation levels were significantly and positively correlated with expression levels, were highly expressed in renal cell carcinoma and lung adenocarcinoma, and activated the mTOR pathway to inhibit cellular autophagy, thereby promoting tumour development [114] (Table 2).

In summary, gene body methylation, similar to promoter DNA methylation, has been shown to play an important role in the occurrence and development of tumours. However, unlike promoter DNA methylation, which mediates tumour development by suppressing oncogene expression, gene body methylation plays a complex role, and the relevant research results are relatively one-sided. As a result, the role of gene body methylation in different tumours still needs to be analysed in specific genes. It is expected that with the gradual exploration of gene body methylation, there will be a complete and accurate theory to explain its abnormal distribution and mechanism of action in tumours.

Gene body methylation and tumour therapy

There are two main types of DNMT inhibitors. One is a nucleotide analogue that inhibits its activity by replacing cytosine during DNA replication and forming a covalent bond with DNMT. Currently, nucleotide analogues of DNMT inhibitors are approved for use by the U.S. Food and Drug Administration (FDA) and show great clinical value in haematologic and solid tumours [115, 116]. Previously, DNA methylation inhibitors were used to reactivate the promoters of methylation-silenced tumour suppressor genes, but these drugs also have a demethylating effect on gene body sequences. Due to the positive correlation between gene body methylation levels and expression levels, gene body demethylation of oncogenes and aberrant metabolic genes in tumours may mitigate their overexpression and thus attenuate their positive effects on tumour development. For example, DNA demethylation agents acting on the gene body of the oncogene C-MYC have been found to have a potential role in attenuating its oncogenic effects [55]. In addition, treatment of cells with 5-aza-CdR revealed that the duration of gene body demethylation nearly overlapped with the time of significant inhibition of cell growth, suggesting that changes in DNA methylation may determine the rate of recovery of cell growth. Genes downregulated after demethylation were upregulated with subsequent DNA remethylation, which was dependent on DNMT3B and local chromatin structure. The combination of DNA methylation inhibitors with DNMT3B inhibitors is therefore expected to prolong the antitumour effect and provides an important direction for future research [55].

Another DNMT inhibitor, which is still in the research stage, is a non-nucleotide analogue that binds directly to the region on DNMT where its methylation acts to inhibit its function [117]. Recent studies have found that 18β-glycyrrhetinic acid (GRA) has the same demethylating effect as DNMT inhibitors in gastric cancer tissues. The expression level of ATP4a, a differentiation-related gene, was significantly downregulated in gastric cancer tissues, and the methylation level of the gene body was negatively correlated with the expression of ATP4a. GRA had a demethylation effect on ATP4a in vivo and in vitro, activating the expression of ATP4a and inhibiting the occurrence of gastric cancer, suggesting that GRA is an effective demethylating agent with certain clinical application prospects for the treatment of gastric cancer [118].

DNMT inhibitors, as widely used antitumor drugs at present, have not received much attention for their demethylation effects on gene body regions. The above studies have provided clues to the role of DNMT inhibitors in gene body methylation, which may inhibit abnormally

Table 2 Gene body methylation and tumorigenesis

Type of cancer	Gene	Functions	References
Thyroid cancer	METTL7A	Metabolic abnormalities	[108]
Hepatocarcinoma	HBx	Dysregulation of gene expression	[109]
Lymphoma	BCL6	Oncogene overexpression	[110]
Testicular cancer	miR-199a	Promotes tumour invasion	[111]
Lung cancer	ITPKA	Promotes tumour invasion	[112]
Renal cell carcinoma and lung adenocarcinoma	CARD11	Inhibits autophagy	[114]

Wang *et al. Clinical Epigenetics*　　(2022) 14:154

Page 9 of 14

overexpressed genes in tumours and have a certain therapeutic effect on tumours. Therefore, future studies can investigate gene body methylation levels as a new therapeutic target, focussing on the role and effects of DNMT inhibitors on gene body methylation and identifying combination drugs or antitumor drugs that specifically target gene body methylation.

Gene body methylation and prognostic assessment of tumours

GenE body methylation abnormalities, as a cause of tumorigenesis, are of great significance as prognostic biomarkers in tumours. There is currently widespread interest in comparing the impact of methylation differences on survival through the analysis of biological information and finding corresponding biomarkers.

In head and neck squamous cell carcinoma, lincRNA C5orf66-AS1 methylation levels are significantly negatively correlated with overall survival, which is a promising prognostic biomarker in head and neck squamous cell carcinoma and can be used to identify HPV-negative patients at risk of recurrence and metastasis [119]. In glioblastoma, the ZMIZ1 gene is a driver of tumour cell migration. ZMIZ1 gene body hypermethylation inhibits tumour cell migration by regulating variable splicing and reducing the expression level of transcripts with higher potency, so patients with ZMIZI gene body hypermethylation tend to have a better prognosis [120]. In lung cancer, 260 exons were found to have differential methylation levels and expression levels in lung cancer tissue compared to normal lung tissue. Among these markers, exon 10 methylation markers were able to significantly differentiate between high-risk and low-risk patients, with promising clinical applications in the prognosis and surgical management of lung cancer [63]. In myelodysplastic syndromes, 57% of patients had DNMT3A aberrant gene body hypomethylation. The multifactorial analysis confirmed that DNMT3A methylation status was an independent prognostic factor in patients with MDS [121].

In gastrointestinal tumours, Liu et al. identified methylation changes in the GFRA1, SRF and ZNF382 genes as potential synergistic biomarkers for predicting gastric cancer metastasis [122]. Debernardi et al. found that the long interspersed nuclear elements 1, located in the second intron of MET oncogene (LINE1-MET), hypomethylation is an important feature of early-onset colorectal cancer (CRC), that LINE1-MET hypomethylation in colorectal adenomas is associated with a high risk of developing colorectal cancer, and that LINE1-MET methylation levels are of great significance for prognostic assessment of patients after surgical resection of polyps [104]. Triple repair nucleic acid exonuclease 2 (TREX2) has a DNA repair function and may exert a tumour suppressive effect. Weigel et al. found that TREX2 gene body hypomethylation resulted in enhancer activity of the TREX2 motif, which promoted TREX2 expression and thus exerted a tumour suppressive effect. TREX2 hypomethylation was strongly associated with prolonged overall survival in laryngeal and colorectal cancers [123]. Bormann et al. found that amphiregulin (AREG) expression was positively correlated with increased progression-free survival and overall survival. Further studies found that AREG gene body methylation levels were significantly negatively correlated with AREG expression levels and therefore may also have some relevance to response after epidermal growth factor receptor (EGFR) therapy. However, a retrospective study found that AREG gene body methylation levels are not as predictive as AREG mRNA levels and cannot yet be used as a clinically reliable biomarker, but they still provide an important clue for further research [124]. In addition, in patients with cholangiocarcinoma, patients with higher methylation levels of PITX2-adjacent noncoding RNAs had relatively longer survival and could be an independent prognostic factor [125] (Table 3).

The above studies show that gene body methylation provides a clear indication of prognosis for a variety of tumours and can be an important supplement for tumour prognosis judgement. Unfortunately, the relationship

Table 3 Gene body methylation and prognostic assessment of tumours

Tumour type	Gene	Correlation of gene body methylation levels and survival	References
Head and neck squamous cell carcinoma	LincRNA C5orf66-AS1	Negative correlation	[119]
Glioblastoma	ZMIZI	Positive correlation	[120]
MDS	DNMT3A	Positive correlation	[121]
Colorectal cancer	LINE-1	Negative correlation	[104]
	AREG		[124]
Laryngeal carcinoma	TREX2	Negative correlation	[123]
Cholangiocarcinoma	PITX2	Positive correlation	[125]

Wang *et al. Clinical Epigenetics*   (2022) 14:154

Page 10 of 14

between gene body methylation and tumour prognosis is still at the level of bioinformatics analysis, and its specific mechanism has not been fully explained, which makes gene body methylation not convincing enough as a biomarker and cannot be widely used in different tumours. Future research needs to focus more on exploring the mechanism of gene body methylation and tumour invasion and development, seeking clearer and more definite pathophysiological mechanisms, and translating them into clinical applications to provide a more accurate judgement for tumour treatment and prognosis.

Conclusion and perspective

The study of genome-wide methylation has led to the initial elucidation of the mechanism of gene body methylation occurrence and the regulation of gene expression. However, the specific role of gene body methylation in gene transcription is still widely debated. Gene body methylation suppresses gene expression by promoting chromatin densification [47] and interacting with functional elements in transcribed regions, such as alternative promoters, enhancers, TF binding sites, and repetitive elements [55]. However, numerous studies have shown that gene body methylation plays a positive role in transcription by suppressing spurious gene transcription, regulating alternative splicing, ensuring correct splicing and translation, and ensuring stable and orderly transcription. In addition, gene body methylation regulates transcription, while it is regulated by active transcription. Therefore, the generation of gene body methylation patterns and the specific molecular mechanisms of biological effects are still key questions that need to be solved and are future breakthrough directions in the field of gene body methylation.

In the field of oncology, current studies have shown that gene body methylation is essential for maintaining the stability of gene transcription and genomes. Tumour genomic hypomethylation can reduce chromosomal stability, disrupt genomic imprinting, and lead to abnormal tumour metabolism and differentiation, which are closely related to tumour occurrence, treatment, and prognosis. However, the mechanism of gene body hypomethylation in tumours is still unclear, and there are mainly two conjectures: (1) one conjecture suggests that tumour gene body hypomethylation is an active demethylation process, and TET family proteins, AICDA, and ELP3 are associated with the demethylation of methylated cytosines [126–128]. However, it was found that none of these genes were significantly higher in breast cancer cells than in normal breast epithelial cells [129]. (2) Another hypothesis holds that DNA methylation is passively lost during tumour replication and occurs mainly in late-replicating regions of the genome [130]. However, it has been shown that maladjustment of DNA repair mechanisms or DNA methyltransferases is not the main cause [131], so the cause of partial loss of tumour DNA methylation levels still needs to be further experimentally explored. In addition, Yen et al. found that gene bodies were completely methylated in two choriocarcinoma cell lines but hypomethylated in mesenchymal cells and placental cells [132], suggesting that tumour gene body hypomethylation is also not absolute and that more complex mechanisms may exist.

Current research on promoter DNA methylation is still just the tip of the DNA methylation iceberg. With the iterative update of DNA methylation sequencing technology and the decreasing cost of sequencing, research on DNA methylation will boom. The discovery of oxidative bisulfite sequencing (oxBS-Seq) in recent years has solved the previous problem of being unable to distinguish DNA methylation from hydroxymethylation, making the study of DNA methylation more accurate. With the availability of large amounts of sequencing data, the challenge ahead is to ensure the depth of sequence reads and the optimization of computational methods. By solving the above challenges, the mechanism and function of DNA methylation can be systematically elaborated and refined, and DNA methylation will represent a new generation of epigenetic biomarkers with unprecedented clinical value.

Wang *et al. Clinical Epigenetics*   (2022) 14:154

Page 11 of 14

Declarations

References

1. Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. Science. 1975;187(4173):226–32.
2. Riggs AD. X inactivation, differentiation, and DNA methylation. Cytogenet Cell Genet. 1975;14(1):9–25.
3. Bird AP. CpG-rich islands and the function of DNA methylation. Nature. 1986;321(6067):209–13.
4. Rauch TA, Wu X, Zhong X, Riggs AD, Pfeifer GP. A human B cell methylome at 100-base pair resolution. Proc Natl Acad Sci U S A. 2009;106(3):671–8.
5. Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, et al. Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature. 2010;466(7303):253–7.
6. Zeng J, Nagrajan HK, Yi SV. Fundamental diversity of human CpG islands at multiple biological levels. Epigenetics. 2014;9(4):483–91.
7. Illingworth R, Kerr A, Desousa D, Jørgensen H, Ellis P, Stalker J, et al. A novel CpG island set identifies tissue-specific methylation at developmental gene loci. PLoS Biol. 2008;6(1): e22.
8. Bewick AJ, Schmitz RJ. Gene body DNA methylation in plants. Curr Opin Plant Biol. 2017;36:103–10.
9. Tran RK, Henikoff JG, Zilberman D, Ditt RF, Jacobsen SE, Henikoff S. DNA methylation profiling identifies CG methylation clusters in Arabidopsis genes. Curr Biol. 2005;15(2):154–9.
10. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SWL, Chen H, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. Cell. 2006;126(6):1189–201.
11. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nat Genet. 2007;39(1):61–9.
12. Zemach A, Zilberman D. Evolution of eukaryotic DNA methylation and the pursuit of safer sex. Curr Biol. 2010;20(17):R780–5.
13. Sarda S, Zeng J, Hunt BG, Yi SV. The evolution of invertebrate gene body methylation. Mol Biol Evol. 2012;29(8):1907–16.
14. Hunt BG, Glastad KM, Yi SV, Goodisman MAD. Patterning and regulatory associations of DNA methylation are mirrored by histone modifications in insects. Genome Biol Evol. 2013;5(3):591–8.
15. Zilberman D. An evolutionary case for functional gene body methylation in plants and animals. Genome Biol. 2017;18(1):87.
16. Jeziorska DM, Murray RJS, De Gobbi M, Gaentzsch R, Garrick D, Ayyub H, et al. DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. Proc Natl Acad Sci U S A. 2017;114(36):E7526–35.
17. Faustino NA, Cooper TA. Pre-mRNA splicing and human disease. Genes Dev. 2003;17(4):419–37.
18. Neri F, Rapelli S, Krepelova A, Incarnato D, Parlato C, Basile G, et al. Intragenic DNA methylation prevents spurious transcription initiation. Nature. 2017;543(7643):72–7.
19. Jones PA, Baylin SB. The epigenomics of cancer. Cell. 2007;128(4):683–92.
20. Ligtenberg MJL, Kuiper RP, Chan TL, Goossens M, Hebeda KM, Voorendt M, et al. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3′ exons of TACSTD1. Nat Genet. 2009;41(1):112–7.

21. Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, et al. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature. 2015;520(7546):243–7.
22. Hashimoto H, Vertino PM, Cheng X. Molecular coupling of DNA methylation and histone methylation. Epigenomics. 2010;2(5):657–69.
23. Ooi SKT, Qiu C, Bernstein E, Li K, Jia D, Yang Z, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature. 2007;448(7154):714–7.
24. Hu J-L, Zhou BO, Zhang R-R, Zhang K-L, Zhou J-Q, Xu G-L. The N-terminus of histone H3 is required for de novo DNA methylation in chromatin. Proc Natl Acad Sci U S A. 2009;106(52):22187–92.
25. Zhang Y, Jurkowska R, Soeroes S, Rajavelu A, Dhayalan A, Bock I, et al. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. Nucleic Acids Res. 2010;38(13):4246–53.
26. Elango N, Yi SV. Functional relevance of CpG island length for regulation of gene expression. Genetics. 2011;187(4):1077–83.
27. Hartl D, Krebs AR, Grand RS, Baubec T, Isbel L, Wirbelauer C, et al. CG dinucleotides enhance promoter activity independent of DNA methylation. Genome Res. 2019;29(4):554–63.
28. Lee S-M, Choi W-Y, Lee J, Kim Y-J. The regulatory mechanisms of intragenic DNA methylation. Epigenomics. 2015;7(4):527–31.
29. Lehnertz B, Ueda Y, Derijck AAHA, Braunschweig U, Perez-Burgos L, Kubicek S, et al. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol. 2003;13(14):1192–200.
30. Fuks F, Hurd PJ, Deplus R, Kouzarides T. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic Acids Res. 2003;31(9):2305–12.
31. Li H, Rauch T, Chen Z-X, Szabó PE, Riggs AD, Pfeifer GP. The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. J Biol Chem. 2006;281(28):19489–500.
32. Chang Y, Sun L, Kokura K, Horton JR, Fukuda M, Espejo A, et al. MPP8 mediates the interactions between DNA methyltransferase Dnmt3a and H3K9 methyltransferase GLP/G9a. Nat Commun. 2011;2:533.
33. Epsztejn-Litman S, Feldman N, Abu-Remaileh M, Shufaro Y, Gerson A, Ueda J, et al. De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. Nat Struct Mol Biol. 2008;15(11):1176–83.
34. Bernatavichute YV, Zhang X, Cokus S, Pellegrini M, Jacobsen SE. Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in Arabidopsis thaliana. PLoS ONE. 2008;3(9): e3156.
35. Stroud H, Greenberg MVC, Feng S, Bernatavichute YV, Jacobsen SE. Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. Cell. 2013;152(1–2):352–64.
36. Benetti R, Gonzalo S, Jaco I, Schotta G, Klatt P, Jenuwein T, et al. Suv4-20h deficiency results in telomere elongation and derepression of telomere recombination. J Cell Biol. 2007;178(6):925–36.
37. Pannetier M, Julien E, Schotta G, Tardat M, Sardet C, Jenuwein T, et al. PR-SET7 and SUV4-20H regulate H4 lysine-20 methylation at imprinting control regions in the mouse. EMBO Rep. 2008;9(10):998–1005.
38. Svobodová Kovaříková A, Legartová S, Krejčí J, Bártová E. H3K9me3 and H4K20me3 represent the epigenetic landscape for 53BP1 binding to DNA lesions. Aging (Albany NY). 2018;10(10):2585–605.
39. Kapoor-Vazirani P, Kagey JD, Powell DR, Vertino PM. Role of hMOF-dependent histone H4 lysine 16 acetylation in the maintenance of TMS1/ASC gene activity. Cancer Res. 2008;68(16):6810–21.
40. Ren W, Fan H, Grimm SA, Kim JJ, Li L, Guo Y, et al. DNMT1 reads heterochromatic H4K20me3 to reinforce LINE-1 DNA methylation. Nat Commun. 2021;12(1):2490.
41. Huang X, Zhang X, Zong L, Gao Q, Zhang C, Wei R, et al. Gene body methylation safeguards ribosomal DNA transcription by preventing PHF6-mediated enrichment of repressive histone mark H4K20me3. J Biol Chem. 2021;297(4): 101195.
42. Wagner EJ, Carpenter PB. Understanding the language of Lys36 methylation at histone H3. Nat Rev Mol Cell Biol. 2012;13(2):115–26.
43. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999;99(3):247–57.

44. Weinberg DN, Papillon-Cavanagh S, Chen H, Yue Y, Chen X, Rajagopalan KN, et al. The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape. Nature. 2019;573(7773):281–6.

45. Brinkman AB, Gu H, Bartels SJJ, Zhang Y, Matarese F, Simmer F, et al. Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. Genome Res. 2012;22(6):1128–38.

46. Estarás C, Fueyo R, Akizu N, Beltrán S, Martínez-Balbás MA. RNA polymerase II progression through H3K27me3-enriched gene bodies requires JMJD3 histone demethylase. Mol Biol Cell. 2013;24(3):351–60.

47. Deaton AM, Webb S, Kerr ARW, Illingworth RS, Guy J, Andrews R, et al. Cell type-specific DNA methylation at intragenic CpG islands in the immune system. Genome Res. 2011;21(7):1074–86.

48. Hellman A, Chess A. Gene body-specific methylation on the active X chromosome. Science. 2007;315(5815):1141–3.

49. Ball MP, Li JB, Gao Y, Lee J-H, LeProust EM, Park I-H, et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol. 2009;27(4):361–8.

50. Aran D, Toperoff G, Rosenberg M, Hellman A. Replication timing-related and gene body-specific methylation of active human genes. Hum Mol Genet. 2011;20(4):670–80.

51. Jjingo D, Conley AB, Yi SV, Lunyak VV, Jordan IK. On the presence and role of human gene-body DNA methylation. Oncotarget. 2012;3(4):462–74.

52. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539):317–30.

53. Jones PA. The DNA methylation paradox. Trends Genet. 1999;15(1):34–7.

54. Scandaglia M, Barco A. Contribution of spurious transcription to intellectual disability disorders. J Med Genet. 2019;56(8):491–8.

55. Yang X, Han H, De Carvalho DD, Lay FD, Jones PA, Liang G. Gene body methylation can alter gene expression and is a therapeutic target in cancer. Cancer Cell. 2014;26(4):577–90.

56. Aure MR, Fleischer T, Bjørklund S, Ankill J, Castro-Mondragon JA, Børresen-Dale A-L, et al. Crosstalk between microRNA expression and DNA methylation drives the hormone-dependent phenotype of breast cancer. Genome Med. 2021;13(1):72.

57. Batsché E, Yi J, Mauger O, Kornobis E, Hopkins B, Hanmer-Lloyd C, et al. CD44 alternative splicing senses intragenic DNA methylation in tumors via direct and indirect mechanisms. Nucleic Acids Res. 2021;49(11):6213–37.

58. Singh S, Narayanan SP, Biswas K, Gupta A, Ahuja N, Yadav S, et al. Intragenic DNA methylation and BORIS-mediated cancer-specific splicing contribute to the Warburg effect. Proc Natl Acad Sci U S A. 2017;114(43):11440–5.

59. Lorincz MC, Dickerson DR, Schmitt M, Groudine M. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. Nat Struct Mol Biol. 2004;11(11):1068–75.

60. Lee DY, Hayes JJ, Pruss D, Wolffe AP. A positive role for histone acetylation in transcription factor access to nucleosomal DNA. Cell. 1993;72(1):73–84.

61. Hong L, Schroth GP, Matthews HR, Yau P, Bradbury EM. Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. J Biol Chem. 1993;268(1):305–14.

62. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet. 2008;40(12):1413–5.

63. Li S, Zhang J, Huang S, He X. Genome-wide analysis reveals that exon methylation facilitates its selective usage in the human transcriptome. Brief Bioinform. 2018;19(5):754–64.

64. Naftelberg S, Schor IE, Ast G, Kornblihtt AR. Regulation of alternative splicing through coupling with transcription and chromatin structure. Annu Rev Biochem. 2015;84:165–98.

65. Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, et al. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature. 2011;479(7371):74–9.

66. Maunakea AK, Chepelev I, Cui K, Zhao K. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. Cell Res. 2013;23(11):1256–69.

67. Iannone C, Valcárcel J. Chromatin's thread to alternative splicing regulation. Chromosoma. 2013;122(6):465–74.

68. Luco RF, Allo M, Schor IE, Kornblihtt AR, Misteli T. Epigenetics in alternative pre-mRNA splicing. Cell. 2011;144(1):16–26.

69. Pradeepa MM, Sutherland HG, Ule J, Grimes GR, Bickmore WA. Psip1/ Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. PLoS Genet. 2012;8(5): e1002717.

70. Yearim A, Gelfman S, Shayevitch R, Melcer S, Glaich O, Mallm J-P, et al. HP1 is involved in regulating the global impact of DNA methylation on alternative splicing. Cell Rep. 2015;10(7):1122–34.

71. Blattler A, Yao L, Witt H, Guo Y, Nicolet CM, Berman BP, et al. Global loss of DNA methylation uncovers intronic enhancers in genes showing expression changes. Genome Biol. 2014;15(9):469.

72. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, et al. The accessible chromatin landscape of the human genome. Nature. 2012;489(7414):75–82.

73. Wiench M, John S, Baek S, Johnson TA, Sung M-H, Escobar T, et al. DNA methylation status predicts cell type-specific enhancer activity. EMBO J. 2011;30(15):3028–39.

74. Charlet J, Duymich CE, Lay FD, Mundbjerg K, Dalsgaard Sørensen K, Liang G, et al. Bivalent regions of cytosine methylation and H3K27 acetylation suggest an active role for DNA methylation at enhancers. Mol Cell. 2016;62(3):422–31.

75. Steinhaus R, Gonzalez T, Seelow D, Robinson PN. Pervasive and CpG-dependent promoter-like characteristics of transcribed enhancers. Nucleic Acids Res. 2020;48(10):5306–17.

76. Montibus B, Cercy J, Bouschet T, Charras A, Maupetit-Méhouas S, Nury D, et al. TET3 controls the expression of the H3K27me3 demethylase Kdm6b during neural commitment. Cell Mol Life Sci. 2021;78(2):757–68.

77. Bell JSK, Vertino PM. Orphan CpG islands define a novel class of highly active enhancers. Epigenetics. 2017;12(6):449–64.

78. Lee S-M, Lee J, Noh K-M, Choi W-Y, Jeon S, Oh GT, et al. Intragenic CpG islands play important roles in bivalent chromatin assembly of developmental genes. Proc Natl Acad Sci U S A. 2017;114(10):E1885–94.

79. Choi W-Y, Hwang J-H, Cho A-N, Lee AJ, Lee J, Jung I, et al. DNA methylation of Intragenic CpG islands are required for differentiation from iPSC to NPC. Stem Cell Rev Rep. 2020;16(6):1316–27.

80. Lee E, Wang J, Yumoto K, Jung Y, Cackowski FC, Decker AM, et al. DNMT1 regulates epithelial-mesenchymal transition and cancer stem cells, which promotes prostate cancer metastasis. Neoplasia. 2016;18(9):553–66.

81. Wong KK. DNMT1: a key drug target in triple-negative breast cancer. Semin Cancer Biol. 2021;72:198–213.

82. Rollins RA, Haghighi F, Edwards JR, Das R, Zhang MQ, Ju J, et al. Large-scale structure of genomic methylation patterns. Genome Res. 2006;16(2):157–63.

83. de Koning APJ, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. PLoS Genet. 2011;7(12): e1002384.

84. Sen SK, Han K, Wang J, Lee J, Wang H, Callinan PA, et al. Human genomic deletions mediated by recombination between Alu elements. Am J Hum Genet. 2006;79(1):41–53.

85. Bodega B, Orlando V. Repetitive elements dynamics in cell identity programming, maintenance and disease. Curr Opin Cell Biol. 2014;31:67–73.

86. Shapiro JA, von Sternberg R. Why repetitive DNA is essential to genome function. Biol Rev Camb Philos Soc. 2005;80(2):227–50.

87. Ross JP, Rand KN, Molloy PL. Hypomethylation of repeated DNA sequences in cancer. Epigenomics. 2010;2(2):245–69.

88. Patchsung M, Settayanon S, Pongpanich M, Mutirangura D, Jintarith P, Mutirangura A. Alu siRNA to increase Alu element methylation and prevent DNA damage. Epigenomics. 2018;10(2):175–85.

89. Wong N, Lam WC, Lai PB, Pang E, Lau WY, Johnson PJ. Hypomethylation of chromosome 1 heterochromatin DNA correlates with q-arm copy gain in human hepatocellular carcinoma. Am J Pathol. 2001;159(2):465–71.

90. Costa G, Barra V, Lentini L, Cilluffo D, Di Leonardo A. DNA demethylation caused by 5-Aza-2'-deoxycytidine induces mitotic alterations and aneuploidy. Oncotarget. 2016;7(4):3726–39.

Wang *et al. Clinical Epigenetics*   (2022) 14:154

Page 13 of 14

91. Kim B-H, Cho N-Y, Shin SH, Kwon H-J, Jang JJ, Kang GH. CpG island hypermethylation and repetitive DNA hypomethylation in premalignant lesion of extrahepatic cholangiocarcinoma. Virchows Arch. 2009;455(4):343–51.

92. Suter CM, Martin DI, Ward RL. Hypomethylation of L1 retrotransposons in colorectal cancer and adjacent normal tissue. Int J Colorectal Dis. 2004;19(2):95–101.

93. Pappalardo XG, Barra V. Losing DNA methylation at repetitive elements and breaking bad. Epigenetics Chromatin. 2021;14(1):25.

94. Ikeda K, Shiraishi K, Eguchi A, Shibata H, Yoshimoto K, Mori T, et al. Long interspersed nucleotide element 1 hypomethylation is associated with poor prognosis of lung adenocarcinoma. Ann Thorac Surg. 2013;96(5):1790–4.

95. Klein Hesselink EN, Zafon C, Villalmanzo N, Iglesias C, van Hemel BM, Klein Hesselink MS, et al. Increased global DNA hypomethylation in distant metastatic and dedifferentiated thyroid cancer. J Clin Endocrinol Metab. 2018;103(2):397–406.

96. Choi SH, Worswick S, Byun H-M, Shear T, Soussa JC, Wolff EM, et al. Changes in DNA methylation of tandem DNA repeats are different from interspersed repeats in cancer. Int J Cancer. 2009;125(3):723–9.

97. El-Ahwany E, Hassan M, Elzallat M, Abdelsalam L, El-Sawy MA-H, Seyam M. Association of Sat-a and Alu methylation status with HCV-induced chronic liver disease and hepatocellular carcinoma. Virus Res. 2022;321:198928.

98. Qu GZ, Grundy PE, Narayan A, Ehrlich M. Frequent hypomethylation in Wilms tumors of pericentromeric DNA in chromosomes 1 and 16. Cancer Genet Cytogenet. 1999;109(1):34–9.

99. Qu G, Dubeau L, Narayan A, Yu MC, Ehrlich M. Satellite DNA hypomethylation vs. overall genomic hypomethylation in ovarian epithelial tumors of different malignant potential. Mutat Res. 1999;423(1–2):91–101.

100. Widschwendter M, Jiang G, Woods C, Müller HM, Fiegl H, Goebel G, et al. DNA hypomethylation and ovarian cancer biology. Cancer Res. 2004;64(13):4472–80.

101. Akers SN, Moysich K, Zhang W, Collamat Lai G, Miller A, Lele S, et al. LINE1 and Alu repetitive element DNA methylation in tumors and white blood cells from epithelial ovarian cancer patients. Gynecol Oncol. 2014;132(2):462–7.

102. Park SY, Seo AN, Jung HY, Gwak JM, Jung N, Cho N-Y, et al. Alu and LINE-1 hypomethylation is associated with HER2 enriched subtype of breast cancer. PLoS ONE. 2014;9(6): e100429.

103. Awada Z, Bouaoun L, Nasr R, Tfayli A, Cuenin C, Akika R, et al. LINE-1 methylation mediates the inverse association between body mass index and breast cancer risk: a pilot study in the Lebanese population. Environ Res. 2021;197: 111094.

104. Debernardi C, Libera L, Berrino E, Sahnane N, Chiaravalli AM, Laudi C, et al. Evaluation of global and intragenic hypomethylation in colorectal adenomas improves patient stratification and colorectal cancer risk prediction. Clin Epigenet. 2021;13(1):154.

105. Baba Y, Yagi T, Sawayama H, Hiyoshi Y, Ishimoto T, Iwatsuki M, et al. Long interspersed element-1 methylation level as a prognostic biomarker in gastrointestinal cancers. Digestion. 2018;97(1):26–30.

106. Martín B, Pappa S, Díez-Villanueva A, Mallona I, Custodio J, Barrero MJ, et al. Tissue and cancer-specific expression of DIEXF is epigenetically mediated by an Alu repeat. Epigenetics. 2020;15(6–7):765–79.

107. Huang KK, Ramnarayanan K, Zhu F, Srivastava S, Xu C, Tan ALK, et al. Genomic and epigenomic profiling of high-risk intestinal metaplasia reveals molecular determinants of progression to gastric cancer. Cancer Cell. 2018;33(1):137-150e5.

108. Zhou S, Shen Y, Zheng M, Wang L, Che R, Hu W, et al. DNA methylation of METTL7A gene body regulates its transcriptional level in thyroid cancer. Oncotarget. 2017;8(21):34652–60.

109. Lee S-M, Lee Y-G, Bae J-B, Choi JK, Tayama C, Hata K, et al. HBx induces hypomethylation of distal intragenic CpG islands required for active expression of developmental regulators. Proc Natl Acad Sci U S A. 2014;111(26):9555–60.

110. Lai AY, Fatemi M, Dhasarathy A, Malone C, Sobol SE, Geigerman C, et al. DNA methylation prevents CTCF-mediated silencing of the oncogene BCL6 in B cell lymphomas. J Exp Med. 2010;207(9):1939–50.

111. Cheung HH, Davis AJ, Lee TL, Pang AL, Nagrani S, Rennert OM, et al. Methylation of an intronic region regulates miR-199a in testicular tumor malignancy. Oncogene. 2011;30(31):3404–15.

112. Wang Y-W, Ma X, Zhang Y-A, Wang M-J, Yatabe Y, Lam S, et al. ITPKA gene body methylation regulates gene expression and serves as an early diagnostic marker in lung and other cancers. J Thorac Oncol. 2016;11(9):1469–81.

113. Wong CH, Li CH, Man Tong JH, Zheng D, He Q, Luo Z, et al. The establishment of CDK9/RNA PolII/H3K4me3/DNA methylation feedback promotes HOTAIR expression by RNA elongation enhancement in cancer. Mol Ther. 2022;30(4):1597–609.

114. McGuire MH, Dasari SK, Yao H, Wen Y, Mangala LS, Bayraktar E, et al. Gene body methylation of the lymphocyte-specific gene results in its overexpression and regulates cancer mTOR signaling. Mol Cancer Res. 2021;19(11):1917–28.

115. Azad N, Zahnow CA, Rudin CM, Baylin SB. The future of epigenetic therapy in solid tumours–lessons from the past. Nat Rev Clin Oncol. 2013;10(5):256–66.

116. Li H, Chiappinelli KB, Guzzetta AA, Easwaran H, Yen R-WC, Vatapalli R, et al. Immune regulation by low doses of the DNA methyltransferase inhibitor 5-azacitidine in common human epithelial cancers. Oncotarget. 2014;5(3):587–98.

117. Brueckner B, Garcia Boy R, Siedlecki P, Musch T, Kliem HC, Zielenkiewicz P, et al. Epigenetic reactivation of tumor suppressor genes by a novel small-molecule inhibitor of human DNA methyltransferases. Cancer Res. 2005;65(14):6305–11.

118. Cao D, Zhao D, Jia Z, Su T, Zhang Y, Wu Y, et al. Reactivation of Atp4a concomitant with intragenic DNA demethylation for cancer inhibition in a gastric cancer model. Life Sci. 2020;242: 117214.

119. Sailer V, Charpentier A, Dietrich J, Vogt TJ, Franzen A, Bootz F, et al. Intragenic DNA methylation of PITX1 and the adjacent long non-coding RNA C5orf66-AS1 are prognostic biomarkers in patients with head and neck squamous cell carcinomas. PLoS ONE. 2018;13(2): e0192742.

120. Mathios D, Hwang T, Xia Y, Phallen J, Rui Y, See AP, et al. Genome-wide investigation of intragenic DNA methylation identifies ZMIZ1 gene as a prognostic marker in glioblastoma and multiple cancer types. Int J Cancer. 2019;145(12):3425–35.

121. Zhang Y-Y, Zhou J-D, Yang D-Q, He P-F, Yao D-M, Qian Z, et al. Intragenic hypomethylation of DNMT3A in patients with myelodysplastic syndrome. Clin Chem Lab Med. 2018;56(3):485–91.

122. Liu Z, Zhang J, Gao Y, Pei L, Zhou J, Gu L, et al. Large-scale characterization of DNA methylation changes in human gastric carcinomas with and without metastasis. Clin Cancer Res. 2014;20(17):4598–612.

123. Weigel C, Chaisaingmongkol J, Assenov Y, Kuhmann C, Winkler V, Santi I, et al. DNA methylation at an enhancer of the three prime repair exonuclease 2 gene (TREX2) is linked to gene expression and survival in laryngeal cancer. Clin Epigenet. 2019;11(1):67.

124. Bormann F, Stinzing S, Tierling S, Morkel M, Markelova MR, Walter J, et al. Epigenetic regulation of Amphiregulin and Epiregulin in colorectal cancer. Int J Cancer. 2019;144(3):569–81.

125. Uhl B, Dietrich D, Branchi V, Semaan A, Schaefer P, Gevensleben H, et al. DNA methylation of PITX2 and PANCR is prognostic for overall survival in patients with resected adenocarcinomas of the biliary tract. PLoS ONE. 2016;11(10): e0165769.

126. Barreto G, Schäfer A, Marhold J, Stach D, Swaminathan SK, Handa V, et al. Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. Nature. 2007;445(7128):671–5.

127. Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature. 2010;463(7284):1042–7.

128. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, et al. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell. 2011;146(1):67–79.

129. Hon GC, Hawkins RD, Caballero OL, Lo C, Lister R, Pelizzola M, et al. Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. Genome Res. 2012;22(2):246–58.

130. Zhang W, Klinkebiel D, Barger CJ, Pandey S, Guda C, Miller A, et al. Global DNA hypomethylation in epithelial ovarian cancer: passive demethylation and association with genomic instability. Cancers (Basel). 2020;12(3):764.

Wang *et al. Clinical Epigenetics*   (2022) 14:154

Page 14 of 14

131. Wild L, Flanagan JM. Genome-wide hypomethylation in cancer may be a passive consequence of transformation. Biochim Biophys Acta. 2010;1806(1):50–7.
132. Yen BL, Hwa H-L, Hsu P-J, Chen P-M, Wang L-T, Jiang S-S, et al. HLA-G expression in human mesenchymal stem cells (MSCs) is related to unique methylation pattern in the proximal promoter as well as gene body DNA. Int J Mol Sci. 2020;21(14):5075.

Publisher's Note