Genomics 104 (2014) 352–357

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Review

Loss of 5-hydroxymethylcytosine in cancer: Cause or consequence?

Gabriella Ficz *, John G. Gribben

Centre for Haemato-Oncology, Barts Cancer Institute, EC1M 6BQ London, UK

ARTICLE INFO

Article history:
Received 21 May 2014
Accepted 22 August 2014
Available online 30 August 2014

Keywords:
5-Hydroxymethylcytosine
DNA methylation
Reprogramming
Epigenetics and cancer

ABSTRACT

Discovery of the enzymatic activity that catalyses oxidation of 5-methylcytosine (5mC) to generate 5-hydroxymethylcytosine (5hmC) mediated by the MLL (KMT2A) fusion partner TET1 has sparked intense research to understand the role this new DNA modification has in cancer. An unambiguous picture has emerged where tumours are depleted of 5hmC compared to corresponding normal tissue, but it is not known whether lack of 5hmC is a cause or a consequence of tumourigenesis. Experimental data reveals a dual tumour-suppressive and oncogenic role for TET proteins. *Tet2* mutations are drivers in haematological malignancies but *Tet1* had an oncogenic role in MLL-rearranged leukaemia, where *Tet1* is overexpressed. Overexpression of *Tet2* in melanoma cells re-established the 5hmC landscape and suppressed cancer progression but inhibiting *Tet1* in non-transformed cells did not initiate cellular transformation. In this review we summarise recent findings that have shaped the current understanding on the role 5hmC plays in cancer.

Contents

1. Introduction . . . 352
2. Distribution and targeting of TET proteins and 5hmC in development . . . 353
3. Modulation of TET activity by cofactors and vitamins . . . 353
4. Mechanism of 5hmC loss in cancer . . . 354
   4.1. Loss of 5hmC in solid cancers . . . 354
   4.2. Loss of 5hmC in haematological malignancies . . . 355
5. 5hmC maintenance through cell division . . . 355
6. Reading 5hmC, 5fC and 5caC (5oxiC) . . . 355
7. Outlook . . . 356
References . . . 356

1. Introduction

DNA methylation became a cancer research focus in the 1970s when analysis of tumours revealed significant aberrations in 5mC distribution, i.e. hypermethylation of gene promoter CpG islands (or CGIs — genomic regions with dense CG content), and gene silencing as a consequence, against a background of genome-wide hypomethylation [1,2]. In somatic cells DNA methylation occurs in CG dinucleotides and about 70–85% of them are constantly methylated, with the exception of CGIs, which tend to be demethylated. Methylation is catalysed by DNA methyltransferases such as DNMT1, the maintenance methyltransferase, which copies methylation to the newly synthesised strand during cell division and DNMT3A and DNMT3B, de novo methyltransferases — enzymes that can methylate DNA in the absence of a hemimethylated template [3]. This heritable epigenetic modification present in all vertebrates is needed for normal development as demonstrated in knockout mice where functional loss of DNA methyltransferases is either early embryonic lethal (*Dnmt1* and *Dnmt3b*) or perinatal lethal (*Dnmt3a*) [4,5]. Due to the high rate of mutation of methylated CG due to spontaneous deamination to generate TG, the frequency of this CG dinucleotide in the genome is significantly reduced (about 4 fold) compared to other dinucleotides such as GC. Although somatically methylated, the genome undergoes global demethylation in early mammalian embryogenesis (during germ cell differentiation and in the zygote), a process associated with totipotency and pluripotency [6–8]. It has been a long-standing

* Corresponding author.
E-mail address: g.ficz@qmul.ac.uk (G. Ficz).

http://dx.doi.org/10.1016/j.ygeno.2014.08.017
0888-7543/

question how DNA demethylation occurs and what enzyme catalyses the break of the C – C bond between the cytosine ring and the methyl group, often with claims of activities that proved to be difficult to validate independently [9].

One of the hypotheses for DNA demethylation involved oxidation of 5mC to 5hmC. 5hmC as a modified base had been previously identified in mammalian tissues but ignored and forgotten largely due to controversy [10]. In 2009, the presence of 5hmC was proven unequivocally and the enzymatic process converting 5mC to 5hmC was identified [11,12]. Compared to the relatively constant levels of 5mC in somatic tissues (3–4% of total cytosines), 5hmC levels are significantly lower and vary greatly depending on the cell type (0.1%–0.7% of all cytosines) [13,14]. The Tet family of Fe(II)- and α-ketoglutarate (α-KG) dependent dioxygenases comprises three proteins (TET1, TET2, TET3) each with varying abundance in different cell types. *Tet1* and *Tet2* are relatively highly expressed in embryonic stem cells and early embryogenesis and their levels drop as cells exit pluripotency and undergo differentiation. Later in development *Tet2* is highly expressed in the haematopoietic system while *Tet3* is overexpressed in germ cells/oocyte, brain tissue and more ubiquitously in somatic cells. TET proteins are responsible for generating all of the 5hmC in the genome [15] and lack of all three TET proteins is incompatible with proper development and such cells contribute poorly to a developing embryo in chimera assays [16]. All three proteins catalyse further oxidation steps generating 5fC and 5caC with varying efficiencies [17,18]. The level of these intermediates is ~20 (5fC) and ~3 (5caC) in 10^6 cytosines per genome (or 0.002% 5fC/C and 0.0003% 5caC/C), at least two orders of magnitude fewer than 5hmC indicating that these modifications could be true demethylation intermediates [18]. Both 5fC and 5caC can be excised by TDG [17,19], thus closing the demethylation cycle [20,21]. Acute downregulation of TDG activity generates up to 7 fold increase in 5fC in mouse embryonic stem (ES) cells, indicating lack of excision of the oxidised 5hmC [22]. It is thought that the main role of 5hmC lies in DNA demethylation but 5hmC is relatively stable and abundant in the genome (see Review by Pfeifer and colleagues, this issue). The question thus arises, if TET proteins are capable of iteratively oxidising 5mC to 5caC then why does this reaction stop at 5hmC so frequently? Current data indicates that perhaps 5hmC and other oxidised forms may have epigenetic roles other than functioning as DNA demethylation intermediates (discussed below). There is a possibility that differences among TET proteins might be reflected by their demethylation potential since TET2 seems to be somewhat more efficient in generating 5fC and 5caC [18]. It is notable that the Tet2 gene region has undergone chromosomal inversion and has lost its CXXC domain during evolution, a chromatin binding protein domain which is retained by both *Tet1* and *Tet3* [23]. The separated region encodes a gene called *Cxxc4* or *IDAX* and has been shown to regulate TET2 protein levels [23]. This is interesting since TET1 and TET2 are targeted to different sites in ES cells, and are thought to have different functions in ES cells [24].

Before the discovery of TET enzymatic activity in 2009, Tet1 was known as the fusion partner of MLL in patients with AML [25,26]. Later it was found that indirect inhibition of TET activity, particularly in AML subclasses with mutations in metabolic enzymes (as discussed below), correlated with promoter hypermethylation. This suggests that lack of TET activity led to increased methylation and subsequent silencing of the majority of genes affected [27,28]. Numerous loss-of-function mutations of *Tet2* have been identified in myeloid cancers where *Tet2* was shown to be a critical tumour suppressor as detailed below. Nevertheless, it has been recently shown that Tet1 plays an oncogenic role in MLL-rearranged leukaemia where the endogenous *Tet1* gene is directly targeted and overexpressed by the MLL fusion proteins [29]. In this study Huang et al. showed in an MLL fusion-induced transformation assay, that downregulation of *Tet1* led to reversal of the tumourigenic phenotype whereas overexpression of *Tet1* exacerbated it. These results suggest contradictory roles for TET proteins in cancer initiation and progression but perhaps such behaviour is more indicative of a mechanism where TET proteins are participating in reinforcing an existing cellular identity and transcriptional network. Nevertheless, it remains to be seen if aberrant targeting of TET proteins could have a significant role in cellular transformation.

2. Distribution and targeting of TET proteins and 5hmC in development

It took over 50 years from the first observation of 5hmC in bacteriophages to its recognition as a new epigenetic modification [30,31]. Compared to 5mC which is relatively abundant in somatic tissues, absolute levels of 5hmC are highly variable both in genome-wide and in specific genomic single base locations and there seems to be no direct correlation between 5mC levels and 5hmC [32] with the exception of complete lack of 5mC where no 5hmC is generated, as detailed below. This is indicative of a highly controlled targeting mechanism of functional TET proteins to the chromatin to generate the 5hmC patterns observed, mechanisms that might be specific to a particular cellular state [33]. Moreover, this control is exerted differentially among the three TET proteins as evidenced by a series of recent observations. In ES cells, association of TET1 to the chromatin is stronger than that of TET2, which could be explained by the lack of the CXXC domain [34]. Also, TET1 and TET2 are targeted to different genomic regions in ES cells, where TET1 is more associated with promoters while TET2 regulates oxidation of gene bodies as shown in Fig. 1 [24]. Interestingly, Huang et al. found that loss of *Tet2* has a more dramatic effect on global 5hmC loss in both *Tet2* shRNA depleted and Tet2 −/− ES cells, even though it is known that the expression level of *Tet2* is lower in ES cells than that of *Tet1* [15]. This study also showed that different regions are affected in shRNA treated ES cells: while *Tet1* shRNA affected mainly promoter regions, loss of *Tet2* resulted in significant loss of 5hmC in gene bodies and at the boundaries of highly expressed exons [24].

One interesting aspect of the abundance of 5mC and 5hmC in adult tissues is that while the majority of somatic cells have relatively constant 5mC levels, cells in the adult stem cell compartment have significantly lower levels for both modifications, therefore it is thought that the presence of 5hmC in healthy tissues is an indication of the differentiation state, with more differentiated cells having more 5hmC and a tissue specific 5hmC level [35,36].

3. Modulation of TET activity by cofactors and vitamins

TET proteins are dioxygenases that use molecular oxygen, Fe(II) and 2-oxoglutarate to oxidise 5mC and convert it to 5hmC. 2-Oxoglutarate (2OG also known as α-ketoglutarate) is a cofactor produced in the Krebs cycle by isocitrate dehydrogenase (IDH1/2/3) through

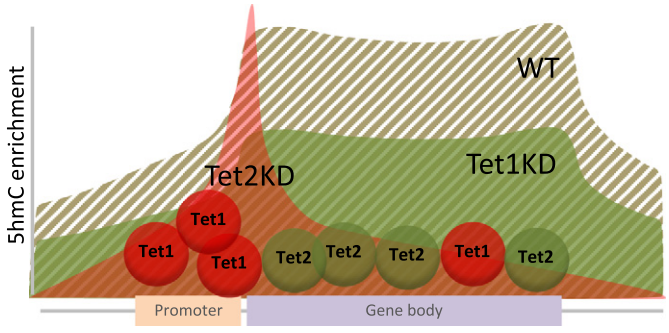

Fig. 1. Distinct distribution patterns for Tet proteins and associated 5hmC loss upon Tet downregulation. The highest density of Tet1 binding sites is located around the gene promoter and loss of Tet1 leads to uniform decrease of 5hmC over the gene body and promoter regions (green relative to the striped area). Loss of Tet2 in mouse embryonic stem cells leads to 5hmC decrease over the gene body indicating specific targeting of Tet2 to gene bodies.