DNA methyltransferases and methyl-binding proteins of mammals

Jie Lan, Song Hua, Xiaoning He, and Yong Zhang*

Key Laboratory of Animal Reproductive Physiology & Embryo Technology, Institution of biotechnology, College of Veterinary Medicine, Northwest A & F University, Yangling 712100, China
*Correspondence address. Tel: +86-29-87080092; Fax: +86-29-87080085; E-mail: zhangyong19561013@yahoo.cn

In mammals, DNA methylation, characterized by the transfer of the methyl group from S-adenosylmethionines to a base (mainly referred to cytosine), acts as a major epigenetic modification. In parallel to DNA sequences arrangement, modification of methylation to DNA sequences has far-reaching influence on biological functions and activities, for it involves controlling gene transcription, regulating chromatin structure, sustaining genome stability and integrity, parental imprinting and X-chromosome inactivation, as well as limiting transposable elements in mammals [2].

DNA methylation preferentially occurs at the C5 position of cytosine in the context of CG, forming the minor bases, 5'-methylcytosines, which account for approximately 1% in the mammalian genome [3] and are mostly found in CpG islands [2]. The so-called CpG island features at least 200 bp in length and a G + C content of 50% as well as a CpG frequency (observed/expected) of 0.6 [4]. The enzymes responsible for DNA methylation are referred to as DNA methyltransferases (DNMTs) that catalyze the reaction through the transfer of the methyl group from S-adenosylmethionine (AdoMet) to cytosine [5]. Murine DNMT1 was the first reported mammalian DNMT by Gruenbaum et al. [6] in 1982. During the past decades, several enzymes harboring methyltransferase activity have been identified.

Although DNA methylation patterns could be established by DNMTs, there emerges a question of how to decipher the information encoded in the methylated sequences. Up to now, methyl-CpG-binding proteins are regarded as the interpreters of the DNA methylation signal, for they have the ability to bind methylated DNA. A family including MeCP2, MBD1, MBD2, MBD3, MBD4, characterized by a methyl-binding motif, has been identified in the mammals [7]. Seemingly, it is indispensable for methyl-CpG-binding proteins to harbor a methyl-binding motif. Nevertheless, it is challenged by the discovery of a new protein named Kaiso without a methyl-binding motif [8].

In this review, a brief overview of the methyltransferases in mammalian will be presented, followed by descriptions of methyl-binding proteins (MBDs).

Mammalian DNA MTases

Here we mainly focus on DNMT1, DNMT2, and DNMT3. They all have 10 conserved motifs: I–X in the carboxyl-terminal region, among which motifs I and X create Adomet-binding site by folding together; motif IV encompasses the prolylcysteinyldipeptide serving as the active
site; motif VI functions in protonating the three positions of the target cytosine via the glutamyl residue within it; and between motifs VIII and IX is the target recognition domain [5]. And their genomic localizations involving four mammals were shown in Table 1.

DNMT1 and its isoforms

Up-to-date, four isoforms have been identified for DNMT1, namely DNMT1s, DNMT1o, DNMT1b, and DNMT1\(^{\Delta E3-6}\) [9,10]. Interestingly, an untranslated transcript, DNMT1p, has also been discovered [11].

DNMT1s (Fig. 1), also known as DNMT1 or DNMT1a, consists of N-terminal regulatory domain and C-terminal catalytic domain connected by glycine–lysine (GK) repeat [9], both of which are indispensable, since the normal activity of DNMT1s is based on the correct folding conformation [12].

As the maintenance MTase, DNMT1s functions in the heritance of pre-existing methylation patterns from parent strands to newly synthesized daughter strands [9]. In fact, maintenance of DNA methylation is not performed \textit{bona fide} with an error frequency of approximately 5% per CpG site per cell division [13], which has been established by quantitative studies of endogenous CpG sites [14]. The overall methylation state, however, stays stable [13]. How does DNMT1s work? It is coupled with replication fork. During the early S phase of cell cycle of somatic cells, DNMT1s, after being translated, is transported into nucleus by virtue of nuclear localization signal (NLS), targets to replication foci under the guide of targeting replication foci (TRF), and then attaches to hemimethylated DNA strands embedded in euchromatin in a complex of DMAP1 and PCNA via NLS-containing domain, zinc-binding domain as well as catalytic domain [15]. When entering into the late S phase, replication foci appear in heterochromatin and the complex becomes HDAC2 attached [16]. In the process, polybromo-1 homology domain (PHD) of DNMT1s may interplay with HDAC, which could facilitate the maintenance of repression state of chromatin [17]. Alternative pathway is replication independent, which occurs during G2 and M phases and could be regarded as an auxiliary way [18].

As for its distribution, DNMT1s is found in embryo pre-implantation stages, although global genomic demethylation begins after fertilization. In mouse, although in low concentration, it was confined to cytoplasm at pronuclear and one-cell stage and distributed wholly in the cell from the two-cell stage [19]. It has been reported that proper

### Table 1: Genomic localizations of DNMTs in four different species of mammals

<table>
<thead>
<tr>
<th>Species</th>
<th>DNMT1</th>
<th>DNMT2</th>
<th>DNMT3a</th>
<th>DNMT3b</th>
<th>DNMT3L</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Homo sapiens</em></td>
<td>19p13.2</td>
<td>10p15.1</td>
<td>2p23</td>
<td>20q11.2</td>
<td>21q22.3</td>
</tr>
<tr>
<td><em>Mus musculus</em></td>
<td>9A3</td>
<td>2A1</td>
<td>12A2–A3</td>
<td>2A2–A3</td>
<td>10C1</td>
</tr>
<tr>
<td><em>Rattus norvegicus</em></td>
<td>8q13</td>
<td>17q12.3</td>
<td>6q14</td>
<td>3q41</td>
<td>20p12</td>
</tr>
<tr>
<td><em>Bos taurus</em></td>
<td>7q15</td>
<td>13</td>
<td>11</td>
<td>13</td>
<td>1</td>
</tr>
</tbody>
</table>

---

**Figure 1** Schematic drawing of the architecture of human DNMT1, DNMT2 and their splicing isoforms with the exception of bovine DNMT2\(\gamma\). PBD, PCNA-binding domain; NLS, nuclear localization signal; TRF, targeting replication foci; ATRX homology, also known as Cys-rich-region, namely zinc-binding motif; PHD, polybromo-1 homology domain; GK repeat, glycine-lysine repeat; the remaining are the six most conserved motifs, with catalytic center located in motif IV. Interaction domains of HDAC1 and HDAC2 are indicated. DNMT2 only consists of C-terminal catalytic domain. DNMT2\(\gamma\) just has 63 amino acids via a premature stop codon.
concentration of DNMT1s is essential for the maintenance of methylation patterns of imprinted genes during the early embryo development [20], for either its absence or overexpression is lethal to embryos [21]. In the adult tissues, DNMT1s shows a high expression as the predominant form [22]. Once it is dysregulated, normal somatic cells get susceptible to malignancy [23]. It has been established that DNMT1s overexpression serves as a hallmark in cancers, such as prostate cancer [24] and endometrioid carcinomas [25]. However, some evidence shows that genomic DNA methylation was hardly impaired in the absence of DNMT1s in the given cancer cells [26]. It can be explained by the point that different cancer cells have a different dependency on DNMT1s expression for the maintenance of DNA methylation and survival [27, 28].

Although DNMT1 is present in the preimplantation embryos in mouse, it is believed that it fails to function and DNMT1o plays a real role instead [29]. As the oocyte-specific isoform, DNMT1o, shorter than DNMT1s due to the loss of the foremost region of the N-terminal, is located in cytoplasm of both oocytes and preimplantation embryos of porcine and mouse with the exception of eight-cell stage when it traffics into nucleus. This behavior of ‘traffic’ is of significance, for its abnormity could give rise to gene ectopic expression [30]. Unexpectedly, it failed to identify oocyte-specific DNMT1 in sheep [31] and bovine [32]. However, surprisingly, a novel isoform of DNMT1, containing 13 additional amino acids encoded by the unspliced exon between exons 12 and 13, was observed in the sheep oocytes, preimplantation embryos and early fetal lineages other than adult tissues. Its absence could lead to embryonic developmental arrest at late morula stage [31].

Another spliced variant of DNMT1, DNMT1b, which contains 16 additional amino acids comparing with DNMT1s and is first identified in human tissues, retains almost all the biological properties of DNMT1s and displays robust activity, although having a low expression level [33, 34], seeming that DNMT1b acts as an assistant of DNMT1s. Distinctions between them require further investigation. The bovine DNMT1b lacks the DMAP1-binding region, expressing lower than DNMT1a and conforming to a tissue-specific mode of regulation [32].

Later work shows a novel variant DNMT1ΔE3-6 with a deletion of part of the DMAP1 interaction domain and PCNA-binding domain corresponding to the exons 3–5 in MT1KO and DKO cell lines, losing the ability to bind PCNA, but it is still a functional MTase in maintaining genomic methylation [10]. With respect to DNMT1p, it is a spliced transcript, regarded as exclusive and untranslatable in pachytene spermatocyte. An identical transcript, however, has been detected in skeletal muscles. More surprisingly, its translated product has also been detected and found to be the same as DNMT1o [35], which strongly challenges the traditional perspective that DNMT1p is an untranslated transcript [11]. The questions of whether it also exists in other tissues and what roles it respectively plays in spermatogenesis and myogenesis need to be answered.

**DNMT2 and its isoforms**

Among MTases, DNMT2 (Fig. 1) is a severely truncated protein, containing only 10 consensus motifs. It has low expression levels in all the tissues examined, and has no detectable MTase activity [36]. *In vitro* experiments showed that it still has the ability to attach to DNA [37]. Later, evidence showed that DNMT2 has weak MTase activity both *in vivo* and *in vitro* [38, 39], suggesting that the N-terminal domain seems unnecessary for MTase activity. Intriguingly, it has been demonstrated that DNMT2 has specificity in methylation of non-CpG dinucleotide in *Drosophila* [40]. Whether similar mechanisms could be extended to mammals remains to be confirmed. Recently, a new discovery was made that human DNMT2 could catalyze RNA methylation [41]. Given that both DNA and RNA can act as substrates and that the affinity of the former to DNMT2 is lower than that of the latter, it is assumed that DNMT2 might be an evolutionary product for MTase transited from a DNA to an RNA target [42]. Additionally, both in testes and in preimplantation embryos a novel isoform is identified of DNMT2, DNMT2-gamma (DNMT2γ), seemingly a nonfunctional protein caused by a premature stop codon [43].

**DNMT3 and its isoforms**

DNMT3 has three members, including DNMT3a, DNMT3b, and DNMT3L. To our knowledge, DNMT3a owns four isoforms (DNMT3a1 to DNMT3a4) and DNMT3b eight isoforms (DNMT3b1 to DNMT3b8) (Fig. 2).

Compared with DNMT1s, DNMT3a and DNMT3b (also respectively known as DNMT3a1 and DNMT3b1) are shorter in length and own the unique structure of tetrapeptide of proline-tryptophan-tryptophan-proline (PWWP) [44]. The loss analysis of PWWP shows its role in directing to chromatin [45], particularly to major satellite regions in pericentric heterochromatin. Notably, the PWWP motif of NDMT3b but not DNMT3a has the ability to combine DNA [46].

By definition, both DNMT3a and DNMT3b are de novo MTases responsible for establishment of methylation patterns of unmethylated DNA strands during development [47]. They have been experimentally demonstrated to be indispensable for embryogenesis [48]. DNMT3b−/− mouse fails to develop to term, and mouse with loss of DNMT3a could just sustain life for a short period after birth [49]. It could be explained that, during mouse early...
embryo development, DNMT3b appears from the formation of blastocyst and disappears after E9.5, then the switch from DNMT3b to DNMT3a occurs [50], suggesting that they are assigned distinct commitments in different stages. For ES cell lines, both single and combinatorial mutation analysis of DNMT3a and DNMT3b confirm that they are required and could compensate each other to some extent [51]. In human, it exhibits an abundant expression level for DNMT3a and very low level for DNMT3b in most tissues other than testis, thyroid, and bone marrow [36]. Overexpression of either DNMT3a or DNMT3b, however, is associated with tumorogenesis or carcinogenesis depending on cancer types [22,52], reflecting to a certain extent that DNMT3a and DNMT3b have their own unique functions or characteristics. It has been shown that DNMT3a fulfills DNA methylation in a distributive manner, whereas DNMT3b does in a mode of process [53] and that DNMT3a is inclined to methylate major satellite in the pericentromeric region, whereas DNMT3b tends to methylate minor satellite in the centromeric region [51]. Furthermore, in vivo studies reveal that the sequences of RCGY and the YCGR are preferentially methylated by DNMT3a and DNMT3b, respectively [54]. More recently, in vitro studies reveal that mouse DNMT3a tends to methylate linker DNA between two nucleosomes in the absence of H1 protein whereas DNMT3b shows the activity of methylating DNA embedded in the nucleosome core region [55]. Additionally, mutation of DNMT3b not DNMT3a has correlation with immunodeficiency, centromeric instability and facial anomalies syndrome [56]. Reversely, it has been demonstrated that DNMT3a not DNMT3b is responsible for the methylation of most imprinted genes in the presence of DNMT3L [57].

Like DNMT1 variants, the isoforms of DNMT3a and DNMT3b due to alternative splicing also play roles in different stages of animal development and their forms vary across species. DNMT3a2, transcribed from an intronic promoter, features the loss of amino acid residues in N-terminal [58]. However, it retains the ability to interact with DNMT3L to achieve regional methylation, preferentially for euchromatin [59]. DNMT3a3 is established in human cells and characterized by the absence of catalytic motifs and MTase activity. Western blot analysis confirms that it is scarce during cell cycle, but is upregulated out of cell cycle in testicular cells and fibroblasts [60]. DNMT3a4 has been isolated from bovine testis. It also lacks MTase activity because of the loss of 67 amino acids in the central coding region [43].

For DNMT3b2, it lacks part of amino acids between regions of PWWP and Cys-rich coincident with exon 10. However, its de novo MTase activity is hardly impaired compared with that of DNMT3b1 [61]. Furthermore, it is detected predominantly in mouse male gonocytes, in which switch from DNMT3b1 to DNMT3b2 occurs [62]. The substitution of DNMT3b2, DNMT1o, and DNMT3a2 in gametes or embryos for their major isoforms DNMT3b1, DNMT1s, and DNMT3a, respectively, raises a question of whether they may conform to the same or similar regulatory mechanism. Besides having the same defect as DNMT3b2, DNMT3b3 also affords the loss of key regions corresponding to exons 22 and 23 [61], becoming unable to methylate DNA. However, it is shown that in in vitro system human DNMT3b3 has almost the same ability to fulfill de novo methylation as DNMT3b1 [63]. Notably, it presumably acts as a regulator for DNA methylation [64]. Together, DNMT3b3 may exert its function either alone less efficiently or in cooperation with other DNMTs more efficiently. The remaining of the isoforms appears to be catalytically inactive due to the disruption of regions coincident with exon 22 or 23 [61]. DNMT3b4 and DNMT3b5 transcripts also have the same deficiency in exon 11 as DNMT3b2 and DNMT3b3. Moreover, the former has no exon 22 and the latter no exon 23. Both DNMT3b4 and DNMT3b5 have a low expression in the analyzed cell lines and normal human tissues except testes.
When translated, both are truncated motifs IX and X due to the distinctive premature stop codons [65]. It has been reported that DNMT3b4 might be a potential contributor to hepatocarcinogenesis, for its upregulation could cause hypomethylation of pericentromeric satellite regions possibly via competing with DNMT3b3 for targeting to the region, during which DNMT3b3 plays an opposite role [66]. However, an opposite result shows that DNMT3b4 is correlated with hypermethylation of promoter region of RASSF1A gene in human lung cancer cells [67]. More studies are needed to elucidate their exact roles. DNMT3b6 transcript was first isolated from ES cells, and it retains exon 11 when compared with DNMT3b3 transcript [65]. A recent study reports the correlation of polymorphism of DNMT3b6 gene with DNA methylation, suggesting that DNMT3b6 may associate with gene transcription [68]. In mouse ES cells, another two transcripts, DNMT3b7 and DNMT3b8, have been discovered. Like DNMT3b6, they keep exon 11 integrated. Yet, they have no exons 22 and 23, respectively [64]. As for human DNMT3b7, it has been established to be an aberrant transcript involved in abnormal DNA methylation profiles of cancer cells [69]. Their translated products require further study. For bovine, four isoforms of DNMT3b have been obtained, corresponding to human DNMT3b1, DNMT3b3, DNMT3b4, and DNMT3b5 [43].

The last member is DNMT3L which is produced during gametogenesis. It just possesses ATRX homology region in the N-terminal, and loses the key catalytic domains in the C-terminal, exhibiting neither DNA-binding ability nor MTase activity. Whereas it serves as a regulator for DNMT3a, DNMT3b, and their isoforms by binding together to induce DNA methylation and improve MTase activity by approximately 1.3–4 folds [70,71]. Moreover, it is indispensable for the establishment of maternally imprinted genes, for its deletion can lead to biallelic expression of normally maternally imprinted genes without disruption of global methylation level [72]. In addition, it can be attached to HDAC to induce de novo methylation via interaction of its zinc finger within the ATRX domain with histone H3 tails, which are unmethylated at Lys4 [73].

Lastly, strictly speaking, DNMT1 and DNMT3 cannot be functionally divided, for they have overlapping functions. As early as 1992, DNMT1 was first reported to possess the de novo methylation activity in the absence of Cys-rich region in vitro [74], suggesting a connection between DNMT1 and DNMT3 and a structural foundation for DNMT1 to possess de novo MTase activities. Given that DNMT3a/3b also have such a structure, Cys-rich region can be considered as a dual controller, not only confining de novo MTase activity of DNMT3a/3b within a regulable range, but also conferring de novo MTase activity on DNMT1 in the given circumstances. Another experiment shows that DNMT1 does play a role both in maintenance and in de novo methylation [75]. Furthermore, a study confirms cooperation among DNMT1, DNMT3a, and DNMT3b in maintenance of DNA methylation [76]. Taken together, their names fail to live up to their functions. It is noteworthy that cooperation among DNMTs also exists in carcinogenesis, although their expressions differ in various cancers [77,78]. The extent to which they are involved in cancer remains to be investigated.

DNA Methyl-binding Proteins

Now that a huge amount of inheritable information is stored in DNA methylation sequences, who act as the interpreters and how to read them out are other big challenges for researchers. Some interpreters have been identified during the past decades, including MeCP1, MeCP2, MBD1, MBD2, MBD3, MBD4, together with Kaiso [79] (Fig. 3). All these proteins except MBD3 have the common property of binding to methylated DNA [80]. In addition, their genomic localizations in four mammals were shown in Table 2. Here are the general descriptions of these proteins.

MeCP1, a protein complex, includes MBD2 and chromatin remodeling complex NuRD/Mi2 containing HDAC1/2, RbAp46/48 and other proteins, in which MBD2 is responsible for preferentially targeting to methylated nucleosomal DNA [81]. Traditionally, it was regarded that MBD1 belonged to the MeCP1 complex [82]. However,...
this point of view seems not to hold water [83]. Hence, more studies are in demand for resolving the controversy.

MeCP2, regarded as the founding member of MBD family proteins, holds two important domains: methyl-binding domain (MBD), in which there exists an Asx-ST motif that can directly interact with DNA and transcriptional-repression domain (TRD) [84]. Evidence shows that both MeCP2 and MBD1/2 have their own unique methylated DNA-binding regions. The methyl-CpG flanked by A/T bases ([A/T] unique methylated DNA-binding regions. The methyl-CpG shows that both MeCP2 and MBD1/2 have their own transcriptional-repression domain (TRD) [84]. Evidence shows that both MeCP2 and MBD1/2 have their own unique methylated DNA-binding regions. The methyl-CpG flanked by A/T bases ([A/T] ≥ 4) is required for targeting of MeCP2, whereas this rule is not suitable for binding of MBD1 or MBD2 [85]. Additionally, MeCP2 is involved in DNA maintenance methylation in the complex with DNMT1 [86]. Through comparison with MeCP1, it shows, on the one hand, MeCP2 binds to methylated DNA more strongly than MeCP1, leading to more stability for gene silencing. On the other hand, densely methylated CpG sites and a single one is needed for targeting of MeCP1 and MeCP2, respectively [83]. It is noteworthy that mutations of MeCP2 are associated with Rett syndrome, which is regarded as one of the most common causes of mental retardation in females and characterized by a progressive neurological impairment [87]. The possible reason is the overexpression of four neural development-related ID genes due to disruption of MeCP2 according to the recent study [88].

MBD1 can be dissected into three key domains: MBD, a Cys-rich domain (CxxC, including CxxC1, CxxC2, and CxxC3), and the C-terminal TRD. MBD and CxxC3 have the ability to bind methylated and unmethylated DNAs respectively, and TRD can actively function at a distance when MBD interacts with DNA during gene inhibition [89]. A dynamic molecular mechanism of interaction of MBD1 with DNA has been advanced [90]. So far, there emerge at least four isoforms: MBD1v1, MBD1v2, MBD1v3, and MBD1v4 [91].

MBD2 closely resembles MBD3 among MBDs and owns an additional 140 amino acid-long N-terminus compared with MBD3. It can assemble into NuRD/Mi2-like complexes and further form MeCP1 with other proteins [92]. It has been demonstrated that MBD2 is involved in the repression of some tumor suppressor genes in tumor development and growth [93]. However, most recently, it is found that MBD2 is the only methyl-binding protein responsible for high expression of human telomerase reverse transcriptase in hTERT-positive cancer lines [94]. Also, MBD2 has been reported to demethylate such methylated gene promoter as SV40 and GL2T, resulting in their activation [95], which suggests that MBD2 may be employed as a potential target for cancer therapy.

MBD3 is special, for it loses the capability of binding to methylated DNA due to insertion of two amino acids, His30 and Phe34, into the MBD domain. However, as a component of NuRD/Mi2 complex, MBD3 has the ability to combine with HDAC1 and MTA2 by MBD domain [80]. Although being unable to directly contact with DNA, the NuRD/Mi2-MBD3 complex still has the ability to participate in gene suppression as an important indirect mediator. One possible way is that NuRD/Mi2-MBD3 complex is recruited to DNA by some specific DNA-binding proteins [96]. Additionally, it has been found that MBD3 is indispensable for proper differentiation of pluripotent stem cells [97]. Similar to MBD2, another surprising study showed that overexpression of MBD3 can also play a role in global demethylation [98], suggesting that MBD3 acts as a bipartite participator both in methylation and in demethylation.

MBD4, also designated as MED1 and identified using a yeast two-hybrid system, is homologous to bacterial DNA repair glycosylases/lyases and responsible for correcting mutations at methylated CpG sites with its endonuclease activity and maintaining genomic stability [99]. Once it is deficient, CpG mutability accumulates, accompanied by the tumorogenesis [100]. Hence, MBD4 can be considered as a potential candidate as a tumor suppressor gene.

The last one is a novel methyl-binding protein, named Kaiso. It has the ability to bind at least two symmetrically methylated CpG sites, preferably in the context of 5′-CGCG-3′ [101]. Furthermore, it can bind non-methylated DNA sequence TNGCAGGA, although having an affinity 1000-fold lower than that to methylated regions [8]. Evidence shows that it is a global repressor of methylated genes and necessary for early embryo development [102]. More details on Kaiso remain to be elucidated.

Collectively, in terms of genomic structure and chromosome location, it could be inferred that MBD2 and MBD3 share a relatively recent ancestor, MBD2/3, found in invertebrate, whereas other MBDs diverge at a relatively distant period. For DNA MBD, it is presumably regarded from a

<table>
<thead>
<tr>
<th>Species</th>
<th>MeCP2</th>
<th>MBD1</th>
<th>MBD2</th>
<th>MBD3</th>
<th>MBD4</th>
<th>Kaiso</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homo sapiens</td>
<td>Xq28</td>
<td>18q21</td>
<td>18q21</td>
<td>19p13.3</td>
<td>3q21–q22</td>
<td>Xq23</td>
</tr>
<tr>
<td>Mus musculus</td>
<td>XA7.3</td>
<td>18E2</td>
<td>18E2</td>
<td>10C1</td>
<td>6E3</td>
<td>XA3.3</td>
</tr>
<tr>
<td>Rattus norvegicus</td>
<td>Xq37</td>
<td>18q12.2</td>
<td>18q12.1</td>
<td>Not placed</td>
<td>4q42</td>
<td>Xq11</td>
</tr>
<tr>
<td>Bos taurus</td>
<td>X</td>
<td>24</td>
<td>Not placed</td>
<td>7</td>
<td>22</td>
<td>Not placed</td>
</tr>
</tbody>
</table>

DNA methyltransferases of mammals

Table 2 Genomic localizations of methyl-binding proteins in four different species of mammals

<table>
<thead>
<tr>
<th>Species</th>
<th>MeCP2</th>
<th>MBD1</th>
<th>MBD2</th>
<th>MBD3</th>
<th>MBD4</th>
<th>Kaiso</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homo sapiens</td>
<td>Xq28</td>
<td>18q21</td>
<td>18q21</td>
<td>19p13.3</td>
<td>3q21–q22</td>
<td>Xq23</td>
</tr>
<tr>
<td>Mus musculus</td>
<td>XA7.3</td>
<td>18E2</td>
<td>18E2</td>
<td>10C1</td>
<td>6E3</td>
<td>XA3.3</td>
</tr>
<tr>
<td>Rattus norvegicus</td>
<td>Xq37</td>
<td>18q12.2</td>
<td>18q12.1</td>
<td>Not placed</td>
<td>4q42</td>
<td>Xq11</td>
</tr>
<tr>
<td>Bos taurus</td>
<td>X</td>
<td>24</td>
<td>Not placed</td>
<td>7</td>
<td>22</td>
<td>Not placed</td>
</tr>
</tbody>
</table>
TAM protein based on their common abilities to bind DNA and co-existence in numerous animal genomes [103]. All of these methyl-binding proteins build a bridge connecting DNA methylation with gene repression, tumor, as well as DNA repair directly or indirectly.

**Perspectives**

How much information can be stored in the epigenetically modified DNA on earth? Although, up-to-date, a substantial body of information has been revealed, we are convinced that our understanding of DNA methylation is just a tip of the iceberg. More DNA MTases, methyl-binding proteins, and their isoforms may be identified in the future among mammalian species and continuously produced possibly by RNA selective splicing under normal and abnormal conditions. Furthermore, their diverse functions remain to be investigated. For instance, a recent study shows that MeCP2 plays roles not only in binding methylated DNA, but also in regulating RNA splicing [104]. Hence, more delicate networks regulating DNA methylation deserve explorations. For the whole genome, given that not all the CpG sites are functional, some interact with methyl-binding proteins, and others just indicate for them like street signs for travelers, still others do nothing. What are the exact recognition mechanisms involving methyl-binding proteins work is still the focuses for future research. Meanwhile, more studies are needed to investigate the relationship among DNMTs, methyl-binding proteins, tumor/cancer, and development to provide more valuable insights into the epigenetic mechanisms during tumorigenesis/carcinogenesis and embryogenesis/ontogenesis. However, one thing is certain: there is still a long way to go for further analysis of diverse functions of DNA methylation regulators and related complicated mechanisms. Fortunately, the rapidly developing DNA methylation-related technologies will make it possible to uncover these mysteries in the future.

**Funding**

This work is funded by the key scientific and technological special program for the culture of disease-resistance transgenic cattle species (2008ZX08007-004), Government of China.

**References**

15. Fatemi M, Herrmann A, Pradhan S and Jeltsch A. The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. J Mol Biol 2001, 309: 1189–1199.
46 Chen TP, Tsujimoto N and Li E. The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. Mol Cell Biol 2004, 24: 9048–9058.


Peddada S, Yasui DH and LaSalle JM. Inhibitors of differentiation (ID1, ID29 ID3 and ID4) genes are neuronal targets of MeCP2 that are elevated in Rett syndrome. Hum Mol Genet 2006, 15: 2003–2014.

Jorgensen HF, Ben-Porath I and Bird AP. Mbd2 is recruited to both methylated and nonmethylated CpGIs via distinct DNA binding domains. Mol Cell Biol 2004, 24: 3387–3395.


