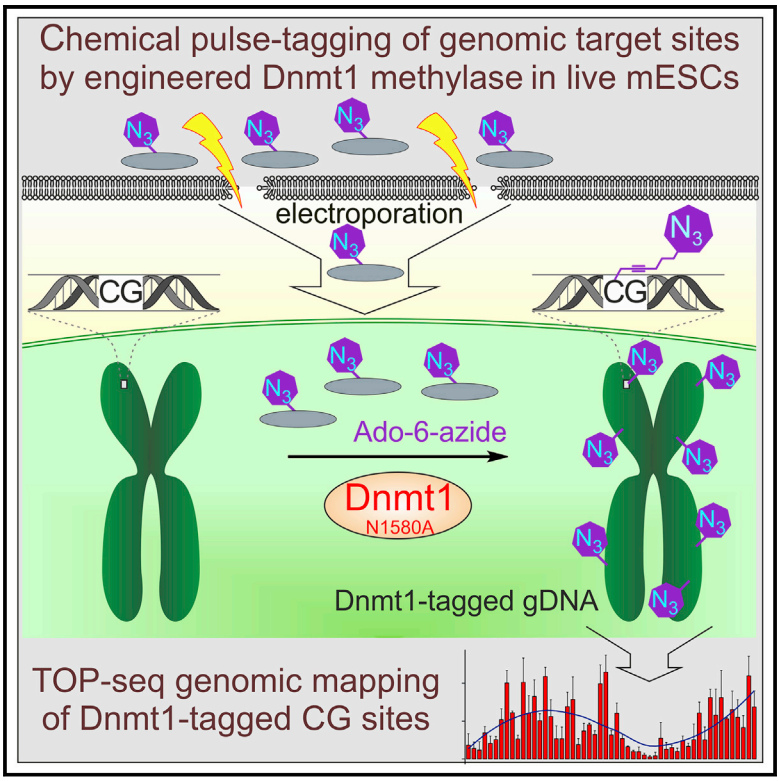
Technology

Molecular Cell

Selective chemical tracking of Dnmt1 catalytic activity in live cells

Graphical abstract



Authors

Vaidotas Stankevičius, Povilas Gibas, Bernadeta Masiulionytė, Liepa Gasiulė, Viktoras Masevičius, Saulius Klimašauskas, Giedrius Vilkaitis

Correspondence

saulius.klimasauskas@bti.vu.lt (S.K.), giedrius.vilkaitis@bti.vu.lt (G.V.)

In brief

Stankevičius et al. report engineering of mouse Dnmt1 methyltransferase for catalytic transfer of extended moieties onto DNA from synthetic cofactor analogs. Genomic installation of the engineered Dnmt1 and pulse internalization of a matching cofactor by electroporation enable selective covalent tagging and precise mapping of catalytic Dnmt1 targets in live pluripotent cells.

Highlights

- A single alanine substitution in Dnmt1 confers catalytic transfer of extended groups
- Electroporation permits facile delivery of AdoMet analogs into live mammalian cells
- Engineered Dnmt1 adds trackable azide tags at its native target sites *in cellulo*
- Dnmt-TOP-seq enables genome-wide tracking of Dnmt1 activity in live mammalian cells

Stankevičius et al., 2022, Molecular Cell *82*, 1053–1065
March 3, 2022 © 2022 The Author(s). Published by Elsevier Inc.
https://doi.org/10.1016/j.molcel.2022.02.008

CellPress

Molecular Cell

CellPress
OPEN ACCESS

Technology

Selective chemical tracking of Dnmt1 catalytic activity in live cells

Vaidotas Stankevičius,[1] Povilas Gibas,[1] Bernadeta Masiulionytė,[1] Liepa Gasiulė,[1] Viktoras Masevičius,[1,2] Saulius Klimašauskas,[1,3,*] and Giedrius Vilkaitis[1,*]

[1]Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius 10257, Lithuania
[2]Institute of Chemistry, Faculty of Chemistry and Geosciences, Vilnius University, Vilnius 03225, Lithuania
[3]Lead contact
*Correspondence: saulius.klimasauskas@bti.vu.lt (S.K.), giedrius.vilkaitis@bti.vu.lt (G.V.)
https://doi.org/10.1016/j.molcel.2022.02.008

SUMMARY

Enzymatic methylation of cytosine to 5-methylcytosine in DNA is a fundamental epigenetic mechanism involved in mammalian development and disease. DNA methylation is brought about by collective action of three AdoMet-dependent DNA methyltransferases, whose catalytic interactions and temporal interplay are poorly understood. We used structure-guided engineering of the Dnmt1 methyltransferase to enable catalytic transfer of azide tags onto DNA from a synthetic cofactor analog, Ado-6-azide, *in vitro*. We then CRISPR-edited the *Dnmt1* locus in mouse embryonic stem cells to install the engineered codon, which, following pulse internalization of the Ado-6-azide cofactor by electroporation, permitted selective azide tagging of Dnmt1-specific genomic targets *in cellulo*. The deposited covalent tags were exploited as "click" handles for reading adjoining sequences and precise genomic mapping of the methylation sites. The proposed approach, Dnmt-TOP-seq, enables high-resolution temporal tracking of the Dnmt1 catalysis in mammalian cells, paving the way to selective studies of other methylation pathways in eukaryotic systems.

INTRODUCTION

Epigenetic regulation of cell phenotype and fate via reversible changes in DNA modification and chromatin structure underlies the mechanisms of development, aging, and disease. One of the fundamental epigenetic mechanisms is methylation of cytosines in CpG dinucleotides to 5-methylcytosine (m^5C), which varies across different genetic loci, cells, and tissues. The DNA methylation toolkit in mammals comprises three catalytically active DNA methyltransferases (DNMT or Dnmt for mouse proteins) (Bestor, 2000; Tajima et al., 2016), which use the *S*-adenosyl-L-methionine (AdoMet) cofactor as the methyl group donor. The first characterized mammalian methylase, DNMT1, preferentially acts on hemimethylated CpG sites (Vilkaitis et al., 2005; Jurkowska and Jeltsch, 2016) (Figure 1A) and is thought to be largely responsible for maintaining pre-existing methylation patterns after DNA replication. The other two major types of mammalian methyltransferases, DNMT3A and DNMT3B, show no such substrate preference and are assigned major roles in methylation of unmodified genomic regions (*de novo* methylation). Loss of the DNMT1 function is directly linked to tumorigenesis and chromosomal instability (Eden et al., 2003; Gaudet et al., 2003), whereas *DNMT3B* mutations cause a severe autosomal disease, called ICF syndrome (Hansen et al., 1999; Okano et al., 1999; Xu et al., 1999). Disruption of each individual *Dnmt* gene in experimental mice leads to a distinct but eventually lethal phenotype, emphasizing the complexity and importance of DNA methylation in mammalian development. On the reverse pathway, DNA methylation can be gradually lost by dilution in the absence of proper maintenance methylation during DNA replication (passive demethylation). DNA methylation can also be actively "erased" via enzymatic oxidation of m^5C by the TET oxygenases to 5-hydroxymethylcytosine (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009) and then to 5-formylcytosine and 5-carboxylcytosine, which are efficiently excised by TDG glycosylase (Denis et al., 2011).

Although DNA methylation is essential for mammalian development, the precise interplay among the three methylating enzymes remains elusive (Lee et al., 2014). Apart from substantial data supporting the classical division of the DNMTs into the maintenance and *de novo* functions (Arand et al., 2012), significant "collaboration" and overlap among the three methylation "writers" has been noted in recent studies (Baubec et al., 2015; Liao et al., 2015; Haggerty et al., 2021). Several methodological and technical bottlenecks grossly hinder further advance in determining the exact location, timing, and role of epigenetic marks deposited by each DNMT. First, typically, an aggregate methylation signal, i.e., the methyl groups collectively deposited by all DNMTs, can be measured by the majority of profiling techniques, which complicates dissection of their individual enzymatic contributions. Second, the roles of individual DNMTs in DNA methylation are often inferred from indirect evidence—the

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

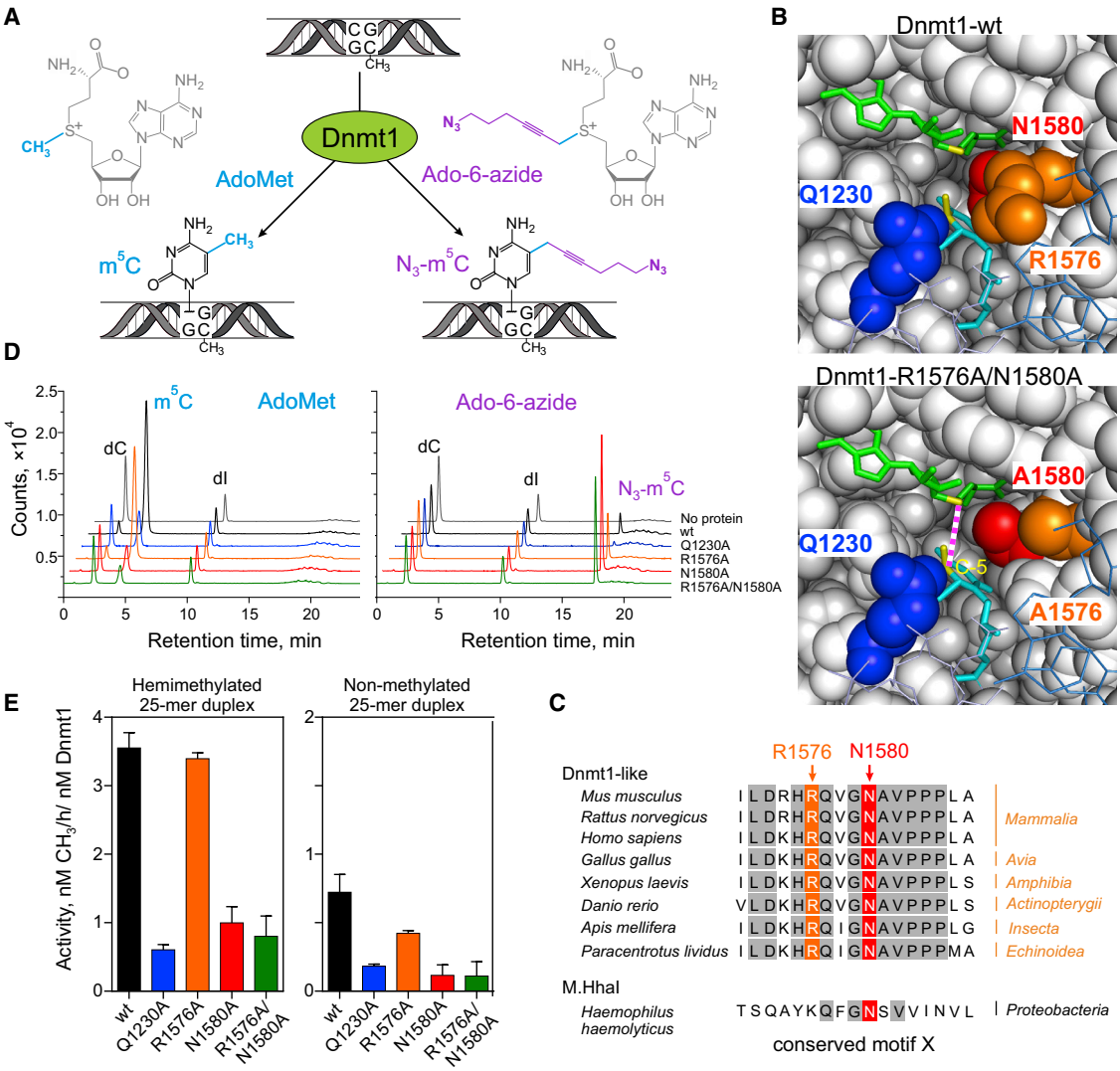


Figure 1. Structure-guided engineering of the Dnmt1 methyltransferase for catalytic transfer of extended functionalized groups

(A) Enzymatic modification of hemimethylated CG/m^{5}CG sites by the mouse Dnmt1 methyltransferase. Biological methylation using the AdoMet cofactor yields m^{5}C, whereas bioorthogonal transfer of an extended side chain with a functional azide group from a synthetic cofactor analog, Ado-6-azide, yields 5-(6-azidohex-2-ynyl)cytosine (N$_3$-m^{5}C) (shown in cyan and magenta, respectively).

(B) Structure of the catalytic/cofactor-binding pocket of WT (PDB: 6w8w) and engineered Dnmt1-R1576/N1580 variants. The methyl-accepting C5 atom of the target 5-fluorocytosine residue (shown in cyan stick presentation) and the methyl-releasing sulfur atom of the bound AdoHcy (green sticks) are connected with a dashed magenta line.

(C) Sequence conservation of the R1576 and N1580 residues in motif X across animal Dnmt1-like proteins and the GCGC-specific bacterial M.HhaI methyltransferase. Identical amino acids are shown in gray boxes.

(D) Enhanced azidoalkylation of DNA by the engineered variants of Dnmt1 methyltransferase. HPLC-MS/MS analysis of nucleosides obtained after hydrolysis of Dnmt1-modified poly(dI-dC)·poly(dI-dC) DNA. Reactions were performed using 6 μM DNA substrate, 100 μM AdoMet or Ado-6-azide, and 400 nM Dnmt1 variant for 40 min at 37 °C.

(E) Methylation activity of Dnmt1 variants toward hemimethylated and nonmethylated 25-mer DNA duplexes. Triplicate reactions containing 6 μM DNA substrate, 5.3 μM [*methyl*-^{3}H]-AdoMet, and 20 nM WT or 100 nM engineered variant as shown were incubated for 40 min at 37 °C. Error bars denote ±SD.

level of gene expression and/or physical localization of the proteins. While constituting an important layer of regulation, the presence of DNMT proteins at specific genomic regions does not always lead to their methylation—as the catalytic activity of all DNMTs is allosterically modulated by numerous external factors (Denis et al., 2011; Li et al., 2011; Jeltsch and Jurkowska, 2016), post-translational modifications (Ling et al., 2004; Estève et al., 2011; Qin et al., 2011), and variations of AdoMet/AdoHcy concentrations, some binding events may serve purely noncatalytic functions. Third, DNMT gene knockouts often used for such studies induce massive undermethylation of the genome, leading to significant phenotypic changes of the cells (Tsumura

et al., 2006; Baubec et al., 2015; Liao et al., 2015; Haggerty et al., 2021). Such dramatic perturbations unavoidably lead to a significantly distorted view of the cellular behavior of the DNMTs.

To circumvent these limitations, we set out to create an experimental platform that permits high-resolution selective tracking of the catalytic contribution of individual DNMT enzymes in live mammalian cells. Our general strategy was thus to replace one of the three DNMTs with a sterically engineered orthogonal variant that preferentially utilizes a synthetically extended AdoMet analog to permit selective chemical tagging of its natural genomic targets (Figure 1A) upon delivery of the cofactor analog inside the engineered mammalian cells. Using the tethered oligonucleotide-primed sequencing (TOP-seq) technique (Staševskij et al., 2017), the deposited chemical groups were then exploited as chemical handles for reading the adjoining sequences and precise genomic mapping of the tagged methylation sites to reveal the Dnmt1-specific genomic methylation profiles. Altogether, we show that our new approach enables selective high-resolution genome-wide tracking of the Dnmt1 catalysis in mouse embryonic stem cells (mESCs), paving the way to deciphering epigenetic mechanisms in numerous developmental and disease model systems.

DESIGN

The creation of an experimental platform for high-resolution selective tracking of the catalytic activity of an individual DNMT enzyme in live mammalian cells entailed the following four key steps: (1) engineering of a bioorthogonal Dnmt-cofactor pair that can deposit trackable chemical groups on DNA, (2) genome editing of a selected mammalian cell line to replace the engineered codon(s) in the *Dnmt1* locus, (3) controlled delivery of the orthogonal cofactor into the live engineered cells, and (4) readout of the genomic positions of the Dnmt-specific chemical marks.

Our previously exploited technology named methyltransferase-directed transfer of activated groups (mTAG) (Dalhoff et al., 2006; Lukinavičius et al., 2007) afforded precise covalent tagging of DNA using appropriate DNA methyltransferases (MTases) as targeting vehicles. During these studies, we found that wild-type (WT) bacterial DNA cytosine-5 MTases were inactive with AdoMet analogs carrying side chains longer than four carbon units, but they could be sterically engineered for acceptance of larger groups (Lukinavičius et al., 2012; Tomkuvienė et al., 2019) by two-three alanine replacements of conserved residues located in the cofactor-binding pocket. Using structure-guided modeling, we selected three potential positions for steric engineering of the cofactor-binding pocket of the mouse Dnmt1 methyltransferase to enable selective deposition of extended chemical moieties containing a terminal azide group from a synthetic cofactor analog, Ado-6-azide. Our selection of a longer, linear propargylic 6-carbon moiety as the tagging unit was motivated by its poor acceptance by most WT MTases (Lukinavičius et al., 2013), which seemed particularly important for confining the transalkylation activity to the engineered MTase in the context of other endogenous AdoMet-dependent MTases

Derivation of a *Dnmt1*^{N1580A} knockin (KI) mESC line was carried out using CRISPR-Cas9 genome editing (Okamoto et al., 2019). To cause minimal alterations of the mESCs in which the activity of Dnmt1 is to be tracked, we decided to make only the required single-codon changes in the genomic copies of *Dnmt1*.

Unfortunately, AdoMet and its analogs are incapable of penetrating the cell membrane and thus are excluded from entering the cells from the medium under standard conditions (McMillan et al., 2005). The permeability of the cell membrane to charged compounds can be increased by installing dedicated membrane transporters (Tucker et al., 2003; Agrimi et al., 2004) or by supplying additional encapsulating delivery compounds (Zawada et al., 2018), which may lead to significant adverse effects. We therefore turned to exploring the suitability of electroporation (Rols, 2006), which uses electric discharge for a temporary enlargement and generation of membrane pores and has proven effective for the introduction of foreign genetic material into mammalian cells.

Finally, genome-wide profiling of the deposited bioorthogonal chemical marks was performed using our previously developed analytical approach, TOP-seq. The approach is based on the "click"-mediated covalent tethering of oligonucleotide primers that permit nonhomologous priming of the DNA polymerase at these internally tagged sites to directly produce adjoining regions for their sequencing and precise genomic mapping (Staševskij et al., 2017)

RESULTS

Engineering Dnmt1 for catalytic transfer of extended moieties from synthetic AdoMet analogs

Guided by available X-ray structures (Song et al., 2012; Adam et al., 2020), we selected three positions for individual and double alanine replacements in Dnmt1 (Figures 1B and 1C). For *in vitro* studies of the transalkylation activity, the WT and four engineered variants were cloned and expressed as truncated (residues 291–1,620) versions in methylotrophic yeast *Pichia pastoris* in high yield and in a predominantly soluble form (Figures S1A and S1B).

Initially, the produced Dnmt1 variants were analyzed by comparing their activity with AdoMet and one of its extensively used extended analogs, Ado-6-azide, which contains the 6-azidohex-2-ynyl side chain carrying a terminal azide group ($-CH_2-C\equiv C-(CH_2)_3-N_3$) (Figure 1A). For kinetic activity assays, we employed two types of previously proven Dnmt1 substrates: 25-mer oligonucleotide duplexes, containing a centrally positioned hemimethylated or unmethylated CpG site and homopolymeric DNAs poly(dI-dC)·poly(dI-dC), or poly(dG-dC)·poly(dG-dC), containing iterate CpI or CpG target sites. In control experiments, the WT Dnmt1 variant isolated from *Pichia pastoris* showed identical specific activity on homopolymeric substrates as previously described baculovirus-derived preparations (Vilkaitis et al., 2005) but hardly detectable activity with the Ado-6-azide cofactor (Figures 1D, 1E, and S1C). Thorough kinetic studies of the engineered Dnmt1 variants under steady-state and single-turnover conditions showed their decreased methylation capacity, which pointed to notable alterations of both AdoMet binding and catalysis (Figures S2 and S3; Tables 1 and S1). To our delight, the transalkylation activity with Ado-6-azide was markedly

Table 1. Single-turnover kinetic parameters of Dnmt1 variants on 25-mer duplex CG-HM DNA for AdoMet-dependent methylation and Ado-6-azide-dependent azidoalkylation

Dnmt1 variant	k_{chem}^{AdoMet} methylation		$k_{chem}^{Ado\text{-}6\text{-}azide}$ azidoalkylation		Cofactor selectivity
	min^{-1}	Mut/WT, fold[a]	min^{-1}	Mut/WT, fold[a]	Ado-6-azide/AdoMet, fold
WT	3.0 ± 0.13	1	0.005 ± 0.005	1	1/600
N1580A	0.104 ± 0.002	1/29	1.4 ± 0.13	265	14
R1576A/N1580A	0.035 ± 0.002	1/86	0.27 ± 0.025	53	8

See Figures S2 and S3 for details.
[a]Values relative to those of WT.

enhanced in the case of Dnmt1-N1580A and R1576A/N1580A variants. The identity of the enzymatically produced "extended" nucleoside, 5-(6-azidohex-2-ynyl)-2′-deoxycytidine (N_3-m^5C), was confirmed by tandem MS/MS (Figures 1D and S1D) and chromatographic comparison with corresponding products of engineered prokaryotic C5-MTases (not shown). Altogether, the enhanced transalkylation power and reduced affinity toward the natural cofactor conferred an inverse cofactor selectivity by the N1580A and R1576A/N1580A variants (14-fold and 8-fold Ado-6-azide over AdoMet preference under single-turnover conditions, respectively; Table 1) and warranted selection of these designer "alkyltransferases" for further comparative *in vitro* studies.

The capacity of the sterically engineered cofactor pocket was further examined against a selection of extended cofactor analogs (Figure 2A). Remarkably, both selected Dnmt1s were highly efficient in transferring propargylic linkers carrying a variety of functional or reporter groups, including a bulky biotin moiety. The catalytic rate of Dnmt1-N1580A with Ado-6-azide cofactor under single-turnover conditions matched within error that of the WT enzyme in the presence of AdoMet (Figure S3; Table 1) and that nearly complete modification of a short hemimethylated duplex could be achieved by brief incubation at cofactor concentrations as low as 1 μM (Figure 2B). Along with efficient two-step labeling of hemimethylated GCG sites in plasmid DNA (Figure S4), this illustrates an immense potential of this new chemoenzymatic tool for orthogonal derivatization and manipulations of DNA.

However, the key factor for achieving the ultimate goal of chemical tracking of Dnmt1 action in living cells is robust transalkylation activity in the presence of the endogenous cofactor. Therefore, we compared the incorporation of azidohexynyl groups from Ado-6-azide in the presence of AdoMet by the selected Dnmt1 variants using different experimental setups. Enzymatic methylation and alkylation of DNA was probed using HPLC-MS/MS analysis of nucleosides (Figures 3A and S5A) or by radiography of electrophoretically separated 5′-^{32}P-labeled strands (Figure 3B). The latter experiment showed that in the presence of equal amounts of both cofactors, the extended cofactor was preferentially used by both engineered variants (the level of alkylation 65%–67%). In contrast, while some alkylation can be observed with Ado-6-azide alone, the WT Dnmt1 yielded no detectable alkylation products in the presence of competing AdoMet. In mouse cell lysates, the exogenous Dnmt1 proteins again showed an inverse catalytic proficiency order for the methylation (WT >> N1580A, R1576A/N1580A, Figure S5B) as compared with the azidoalkylation (WT << R1576A/N1580A < N1580A, Figure S5C) reactions. These observations reassured that the catalytic transfer of bioorthogonal groups by Dnmt1 could be achieved in the background of endogenous AdoMet and other components present inside the mammalian cells.

Engineering mESCs for Dnmt1-directed deposition of extended azide tags

Derivation of a *Dnmt1*^{N1580A} KI mouse embryonic stem E14TG2a cell line was carried out by replacing the AAT codon to GCT in exon 38 using CRISPR-Cas9 genome editing (Okamoto et al., 2019). A 70-nt single-stranded donor template with an additional intronic G to T mutation at the PAM sequence was used to prevent recutting of the KI locus (Figures S6A–S6C). The correctness of the biallelic N1580A substitution was confirmed via restriction analysis and genomic sequencing of adjacent regions. qRT-PCR analysis confirmed that the expression level of Dnmt1-N1580A in KI cells was identical within error with that of Dnmt1-WT in the parental E14TG2a cell line (Figure S6D). The produced KI line, along with the WT and *Dnmt1* KO controls, was then examined for global m^5C levels using HPLC-MS/MS analysis (Figure S6E). The cells were cultured in media with serum and leukemia inhibitory factor (LIF), which exhibit global DNA hypermethylation (Leitch et al., 2013) (Figure 4A). Notably, the observed intermediate genomic m^5C levels in the KI cells indicated that the endogenously expressed Dnmt1-N1580A retained partial methylation potential *in vivo*, in line with its methylation activities observed *in vitro*. Addition of the Ado-6-azide cofactor to the KI mouse ESC lysates conferred efficient azidoalkylation of exogenous pΔL2-14 plasmid DNA containing hemimethylated CGC/Gm^{5}CG sites (Gerasimaitė et al., 2009) by the engineered Dnmt1 but not by the WT control (Figure 4B). These observations clearly suggested that selective deposition of extended groups by the endogenously expressed Dnmt1 variant could be achieved, provided that the synthetic cofactor could be made available inside the mouse cells.

To this end, we examined the suitability of electroporation (Rols, 2006), which uses electric discharge for a temporary enlargement and generation of membrane pores and has proven effective for the introduction of foreign genetic material into mammalian cells. After extensive experimental trials, we established electroporation conditions that permitted a clearly detectable pulse labeling of genomic DNA in the KI cells by exogenous Ado-6-azide (Figure 4C). Typically, serum/LIF medium-grown cells were subjected to electroporation and brief exposure

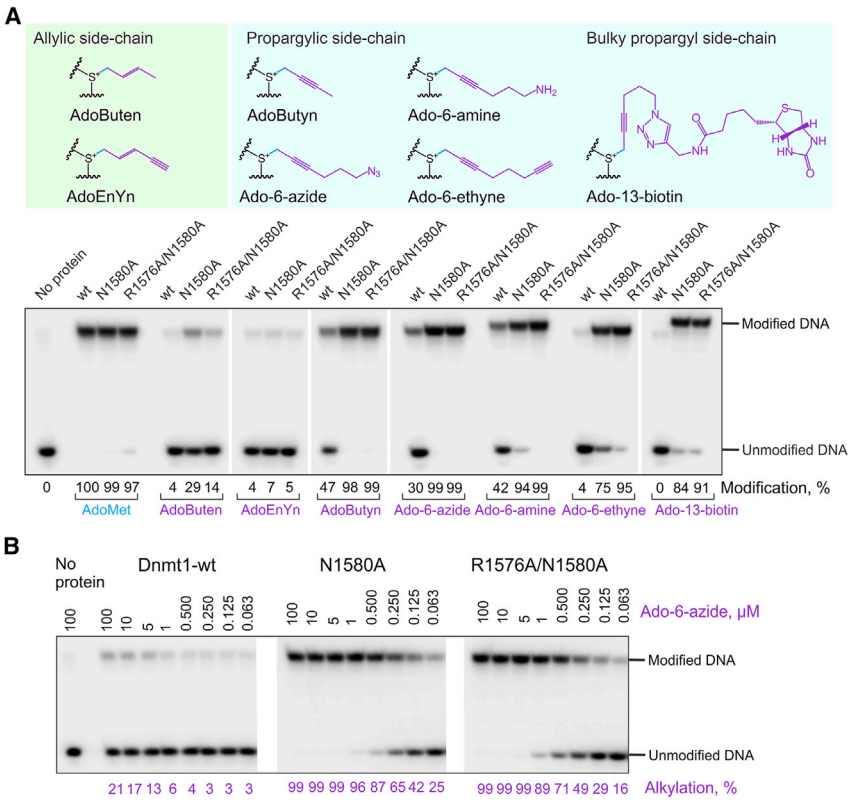


Figure 2. Enhanced catalytic activity of the engineered Dnmt1 variants with AdoMet analogs containing sulfonium-bound propargylic side chains

(A) DNA modification reactions were carried out for 1 h at 37°C using 50 nM hemimethylated 25-mer duplex CG-HM in which the unmethylated strand was 5′-^{32}P-labeled, 100 nM Dnmt1 and 100 μM AdoMet, or its synthetic analog (partial structures shown in top panel). Resulting modified DNA was reannealed with a 125-fold molar excess of an unmodified complementary strand, digested with the methylation-sensitive HhaI endonuclease, fractionated by denaturing PAGE and 5′-^{32}P-labeled strands, and visualized by autoradiography (see also Figure S4).

(B) The effect of Ado-6-azide concentration on activity of Dnmt1 variants under single-turnover conditions. Reactions were carried out using 50 nM DNA, 100 nM Dnmt1, and Ado-6-azide as shown.

(25 min) to serum-free medium containing Ado-6-azide cofactor, and then reversed to serum medium and incubated for an additional 1–6 h. Intragenic incorporation of N_3-m^{5}C in KI mESCs was linearly dose dependent (at least up to 1 mM cofactor) and peaked at ~3 h incubation time (Figures 4D and 4E); our practically attained N_3-m^{5}C levels (with 1 mM Ado-6-azide and 3 h incubation) were about 2 orders of magnitude below the endogenous levels of m^{5}C (0.02% and 4%, respectively; see Figures 4E and S6E). A methyltetrazole dye (MTT) cell viability assay performed 24 h after such electroporation-driven pulse labeling showed a very high survival rate of the WT and KI mESCs both in the presence and absence of exogenous cofactor in the medium (Figure 4F), indicating that neither the electroporation shock nor the genomic incorporation of the extended m^{5}C analog at this level was detrimental to the cells. Altogether, these results clearly demonstrate that a synthetic cofactor can be delivered to and utilized in live mammalian cells by the engineered Dnmt1 in a temporally controlled manner. This provides a new chemical tool for "recording" the genomic footprint of the Dnmt1 catalytic activity in live mammalian cells in the background of native DNA methylation machinery.

Genome-wide selective tracking of Dnmt1 catalysis in live ESCs

Finally, we assessed the suitability of the engineered mESC line to report on the intracellular Dnmt1-dependent methylation events using the Ado-6-azide-dependent chemical pulse-tagging of genomic DNA. We performed genome-wide profiling of the deposited bioorthogonal chemical marks using our previously developed analytical approach, TOP-seq. The approach is based on the "click"-mediated covalent tethering of oligonucleotide primers that permit nonhomologous priming of the DNA polymerase at these internally tagged sites to directly produce adjoining regions for their sequencing and precise genomic mapping (Staševskij et al., 2017). We prepared triplicate Dnmt-TOP-seq libraries (two technical replicates each) from the serum/LIF-cultivated *Dnmt1*^{N1580A} (engineered) and *Dnmt1*^{WT} (control) mESCs following their pulse labeling with 1 mM Ado-6-azide as depicted in Figure 4C. The number of sequencing reads obtained from the Ado-6-azide-treated engineered cells (30–60 M per technical replicate) was 2–3 orders of magnitude higher than that in the controls (Figures S7A and S7B), as were the correlations between biological replicates both genome-wide and within individual genomic elements (Figure S7C). Importantly, we found that the start position of the overwhelming majority (>90%) of mapped *Dnmt1*^{N1580A} reads occurred within 4 nt from a genomic CpG site (Figures 5A and S7D). This was very reassuring, given that Dnmt1 is a CpG-specific methyltransferase, and therefore such reads were attributed as identified CpGs or Dnmt1-modified CpGs (although it is possible that a small fraction of non-CpG reads represent genuine Dnmt1 methylation targets in mESCs, they were excluded from further analysis to minimize noise). The average coverage of identified CpGs was 6–8× per replicate (Figures S7E and S7F) and showed an even distribution in the DNA strands and across the chromosomes (Figure S7G). In contrast, the fraction of CpG-associated reads was only ~10% in the control libraries, indicating their highly random character and thus a negligible contribution from the Dnmt1-WT or *de novo* Dnmt3a or 3b methyltransferases (Figures 5A and S7D).

To establish the suitability of the produced Dnmt1^{N1580A} (further on Dnmt1-TOP-seq) libraries for epigenome studies, we looked at the genome-wide distribution of identified modified

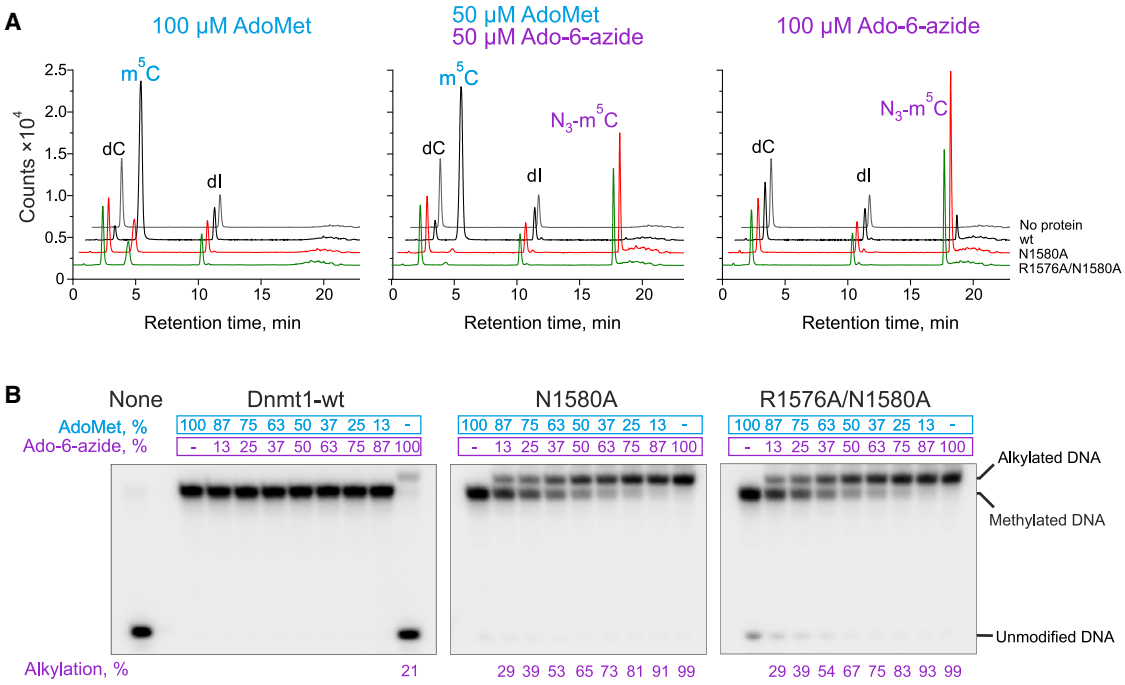


Figure 3. Cofactor selectivity of the wild-type and engineered Dnmt1 variants

(A) Comparative analysis of DNA modification products by the Dnmt1 variants in the presence of AdoMet, Ado-6-azide, and their equimolar mixture under steady-state conditions. DNA modification reactions containing 6 μM poly(dI-dC)·poly(dI-dC) DNA substrate, 400 nM Dnmt1 variants, and 100 μM cofactor were incubated for 40 min at 37°C; modified DNA was hydrolyzed to nucleosides and analyzed using HPLC-MS/MS.

(B) Comparative analysis of DNA modification products by the Dnmt1 variants at varied ratios of the AdoMet and Ado-6-azide cofactors under single-turnover conditions. DNA modification reactions containing 50 nM hemimethylated oligonucleotide, 100 nM Dnmt1, and 100 μM cofactor or their mixture as indicated were incubated for 60 min at 37°C. Reactions were analyzed as described in Figure 2.

CpG sites along a range of genomic features. In protein-coding genes, the CpG modification profile showed a characteristic profile (Jung et al., 2017; Buitrago et al., 2021) with a sharp drop at TSSs followed by a gradual increase toward the 3′ end of genes and a second drop around TTS (Figure 5B). In addition, long regulatory RNAs showed a similar but less pronounced modification profile, in agreement with their gene-like transcriptional status. As a background (untranscribed) reference, we looked at pseudogenes and found that unprocessed pseudogenes indeed maintained a low and uniform level of modification throughout their length. Notably, processed pseudogenes manifested a distinct profile with two DNA modification peaks at TSS and TTS, separated by a valley at gene bodies (Figure 5B). This unexpected observation clearly points at some yet unreported phenomenon that involves epigenetic regulation of these seemingly silent elements (Troskie et al., 2021).

In the next step, we performed functional enrichment analysis (Hansen et al., 2011) of the Dnmt-TOP-seq libraries and identified significant Dnmt1 modification differences between genomic region sets and regulatory elements (Figure 5C). We found strong hypomodification of CpG islands, 5′ UTR/promoter regions, and specific types of noncoding regulatory RNAs, such as miRNAs and snRNAs. In contrast, modified CpGs were enriched in intron/exon coding regions with downstream 3′ UTRs, long noncoding RNAs, and splicing-associated snRNAs. Notably, our analysis revealed contrasting modification of mobile elements (high in SINEs and low in LINEs), in good agreement with previously observed genome-wide methylation of distinct genomic regions in mouse ES-E14TG2a (Zhao et al., 2014). In spite of the generally low methylation levels, the modified positions in CGIs appeared to be well covered in the Dnmt-TOP-seq libraries (Figure S8A). To gain a further system-level understanding of the Dnmt1 activity, we inspected which gene pathways were enriched for Dnmt-TOP-seq modifications. We selected fragments satisfying the criteria of FDR-adjusted p value > 0.05 and absolute log transformed Fisher's estimate >1. We found the majority of such elements to be depleted in modified CpG sites, except for enriched DNA methylation in gene bodies (Figures S8B and S8C). Gene ontology (GO) analysis of the hypomodified genes found that they were strongly enriched for genes related to multicellular organism development processes and cell fate commitment (Figure S8D). Methylated CGIs within promoters showed an enrichment in nuclear division and in pathways related to meiotic cell cycle, gamete generation, and piRNA metabolic processes (Figure 5D), whereas the most highly enriched GO terms (statistical significances are weaker compared with those obtained with promoter CpGs) among genes with methylated CGIs assigned to intergenic CGIs included categories such as development and cell-cell signaling processes (Figure S8E). The specific enrichment of cell cycle and noncoding RNA regulatory functions among methylated promoter-associated CpGs, but not among abundantly methylated

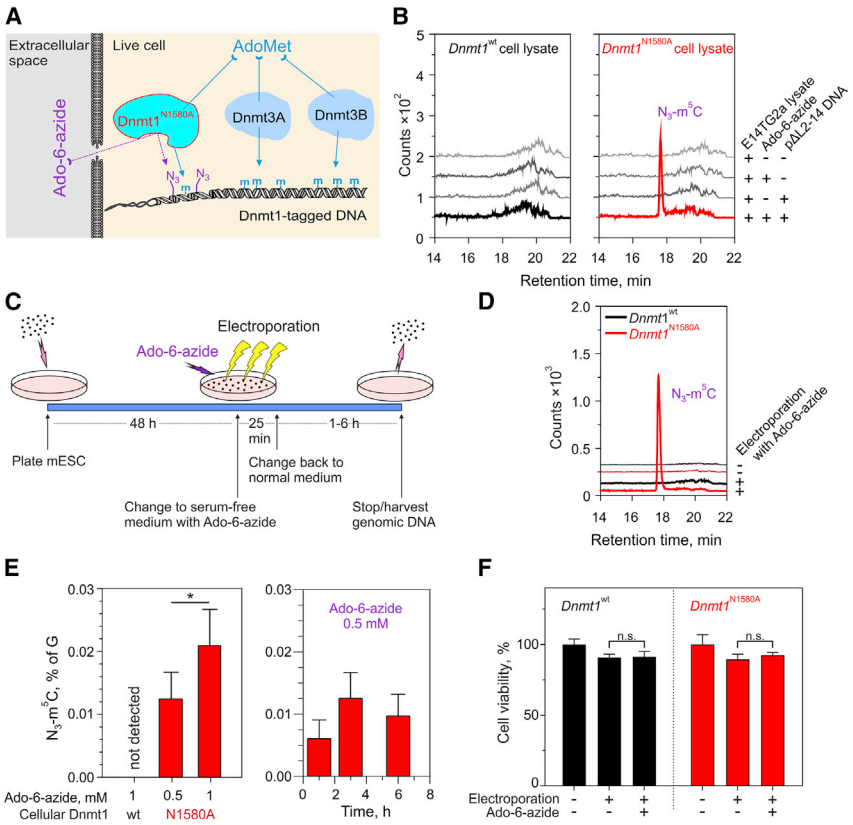


Figure 4. Installation of bioorthogonal functional groups in DNA by endogenous Dnmt1-N1580A in mESC lysates and live cells

(A) Strategy for selective Dnmt1-directed bioorthogonal pulse-tagging of endogenous DNA targets in a living cell. m, methylated cytosine.

(B) Endogenous Dnmt1-N1580A alkyltransferase activity in cell lysate. DNA alkylation reactions were performed for 1 h at 37°C in 20 μL of corresponding mESC lysates supplemented with 100 μM Ado-6-azide and 500 ng of *in vivo*-hemimethylated pΔL2-14 plasmid DNA (Gerasimaitė et al., 2009), and DNA hydrolysates were analyzed for N_3-m^5C using HPLC-MS/MS.

(C) Experimental procedure for selective Dnmt1-directed pulse-tagging of endogenous DNA targets in live mESCs.

(D) HPLC-MS/MS analysis of genomic cytosine modification in *Dnmt1*^{WT} and *Dnmt1*^{N1580A} mESCs after electroporation with 1 mM of Ado-6-azide.

(E) Effects of cofactor concentration and post-electroporation incubation (chase) time on Dnmt1-directed intragenomic incorporation of N_3-m^5C in mESCs.

(F) Viability of WT and engineered mESCs after electroporation in the presence of 1 mM of Ado-6-azide. Cell survival was determined using an MTT assay 24 h after electroporation. Mean ± SD of at least three independent replicates. *p < 0.05; n.s., not significant. See also Figure S6.

genes *per se* or intergenic CGIs, proposed regulatory coherence by this type of CpGs.

In an effort to decouple the contribution of Dnmts in the establishment and maintenance of DNA methylation, we compared the Dnmt1-specific modification profile with genome-wide binding profiles of Dnmt3a and Dnmt3b in mESCs obtained by ChIP-seq (Weinberg et al., 2019). Analysis of intergenic, intragenic, and promoter CGIs revealed a quite intricate interplay of the three DNMTs (Figures 6A, S9A, and S9B), suggesting that methylation of intragenic and promoter CGIs could require contributions from Dnmt1 and Dnmt3b. Remarkably, in LINE and LTR retrotransposons, we can see that the patterns of Dnmt3a1 and Dnmt3b localization show no correlation with the Dnmt1 activity, suggesting that the *de novo* methyltransferases have a weak, if any, impact on the LINE and LTR modification in ESCs (Figure 6B). The latter results are consistent with previous findings implicating Dnmt1 as the key player in silencing of mobile elements and maintaining genome stability (Ran et al., 2013; Min et al., 2020; Haggerty et al., 2021).

Having demonstrated that Dnmt-TOP-seq can produce comprehensive high-resolution genome-wide profiles of the Dnmt1-directed modification, we went on to the level of individual loci to see how well it compares with native m^5C patterns, which represent collective methylation produced by the three cellular Dnmts. A total of 6 selected CGI regions, including 4 hypomodified loci (CGIs of *Tfcp2l1*, *Cdh7*, *Gli2*, and *Ddx18* genes) involved in mESC self-renewal pathways (Das et al., 2013; Ming et al., 2020) or essential for embryonic development (Pieters and van Roy, 2014) and 2 moderately modified regions (CGIs of *Sfi1* and *H1fnt* genes) associated with spermatogenesis (Martianov et al., 2005; Lambrot et al., 2013), were validated using padlock bisulfite sequencing (BS-seq). As expected, BS-seq analysis of the hypomodified regions revealed virtually no CpG methylation, clearly confirming our Dnmt-TOP-seq data and thus excluding the possibility that the lack of Dnmt-TOP-seq reads was due to insufficient sensitivity (Figure S9C). In further agreement, BS-seq analysis revealed 30%–40% methylation coverage in the two moderately modified regions (Figure S9D). Overall, we found a good resemblance between the positional methylation profiles of the latter loci obtained by BS-seq and our method (Figure 6C), except for a notable deviation at the 5′ end of the promoter CGI in *Sfi1* gene. We also performed BS-seq analysis of our generated isogenic *Dnmt1* KO mESCs (fully inactivated Dnmt1). The differential KI-KO BS-seq signal was expected to roughly reflect the Dnmt1-N1580A methylation activity with AdoMet, which in turn could be compared with the Dnmt-TOP-seq profile. The presented data (Figure 6C) indeed show a good concordance (shape similarity) between the differential BS-seq and Dnmt-TOP-seq profiles. The left shoulder of the BS-seq signal in *Sfi1* would be expected to derive from *de novo* Dnmt3 methylation. This is consistent with previously reported DNMT3a recruitment at CGI shores, although individual isoforms may show different cell-specific genomic binding (Manzo et al., 2017). In contrast,

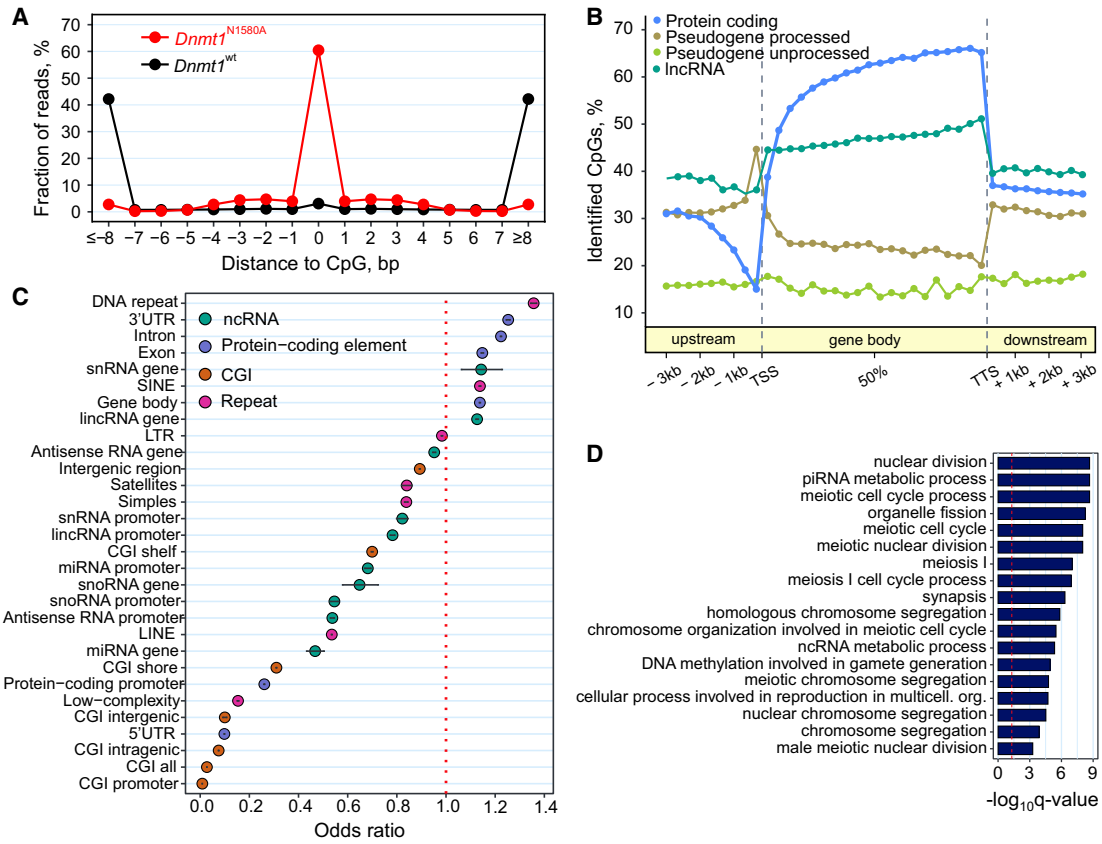


Figure 5. Analysis of genomic Dnmt1 modification sites in mESCs using Dnmt-TOP-seq

(A) Distance distribution of read start positions to a nearest CpG site in the Dnmt-TOP-seq libraries prepared from WT and *Dnmt1*^{N1580A} mESCs 3 h after electroporation with 1 mM Ado-6-azide.

(B) Dnmt-TOP-seq CpG modification profiles along generalized genomic elements for various gene types. Modified CpG sites were computed in the upstream (4 kb from TSS), gene body (from TSS to TTS normalized by gene length), and downstream (4 kb from TTS) regions. Processed pseudogenes are reverse-transcribed copies of mRNAs that lack introns, whereas unprocessed pseudogenes are produced by gene duplication and may contain introns.

(C) Enrichment analysis of Dnmt-TOP-seq genomic elements and regulatory features. Odds ratio denote the enrichment (>1) or the depletion (<1) of particular genomic element terms in the genome-wide DNA modification profile of *Dnmt1*^{N1580} mESC.

(D) Enriched GO terms of genes containing methylated CGI promoters in *Dnmt1*^{N1580A} cells. CGIs bearing at least one modified CpG in all biological replicates were designated for the analysis. q value denotes false discovery rate. See also Figures S7 and S8.

methylation at the central region of intergenic CGI in *H1fnt* was increased in KO cells, suggesting that Dnmt1 could somehow interfere with *de novo* methylation activity at this particular locus in *Dnmt1*^{N1580A} mESCs. Future detailed studies using Dnmt-TOP-seq are needed to gain a better understanding of the mechanisms underlying the temporal activity of Dnmt1 in mESCs.

DISCUSSION

Previously, AdoMet analog-based profiling of protein methylation targets was described in cell lysates (Sohtome and Sodeoka, 2018), and prototype systems were developed for analysis of histone methylation (Wang et al., 2013) and mRNA adenine-*N*6 methylation (Hartstock et al., 2018; Shu et al., 2020) in transfected and/or methionine-deprived cells. Here, we present the first engineered mammalian system that enables activity-based chemical tracking of a specific DNA methylating enzyme in live cells under nearly native conditions. Our first step was the successful engineering of a mammalian Dnmt to an alkynyl transferase that preferentially uses synthetic AdoMet analogs for transfer of bioorthogonal chemical moieties carrying a range of functional and reporter groups on to its genuine targets in DNA (Figures 1 and 2). The transalkylation rate by the engineered Dnmt1 from the Ado-6-azide cofactor is nearly on par with the methyltransfer rate of the WT enzyme in the presence of AdoMet (1.4 and 3.0 min^{-1}, respectively) and is 265-fold higher than that of the WT and Ado-6-azide pair (Table 1). The DNMTs contain well-conserved catalytic domains with ten identified sequence motifs shared between eukaryotic and prokaryotic cytosine-5 MTases. Remarkably, although similar engineering of bacterial orthologs typically requires two replacements in conserved motifs IV and X (Lukinavičius et al., 2012; Deen et al., 2018), the largest enhancement in transalkylation activity and cofactor selectivity was achieved by a single replacement of the conserved Asn1580 to Ala in motif X of Dnmt1 (Figures 1B and 1C). The side chain of this residue serves to sterically confine

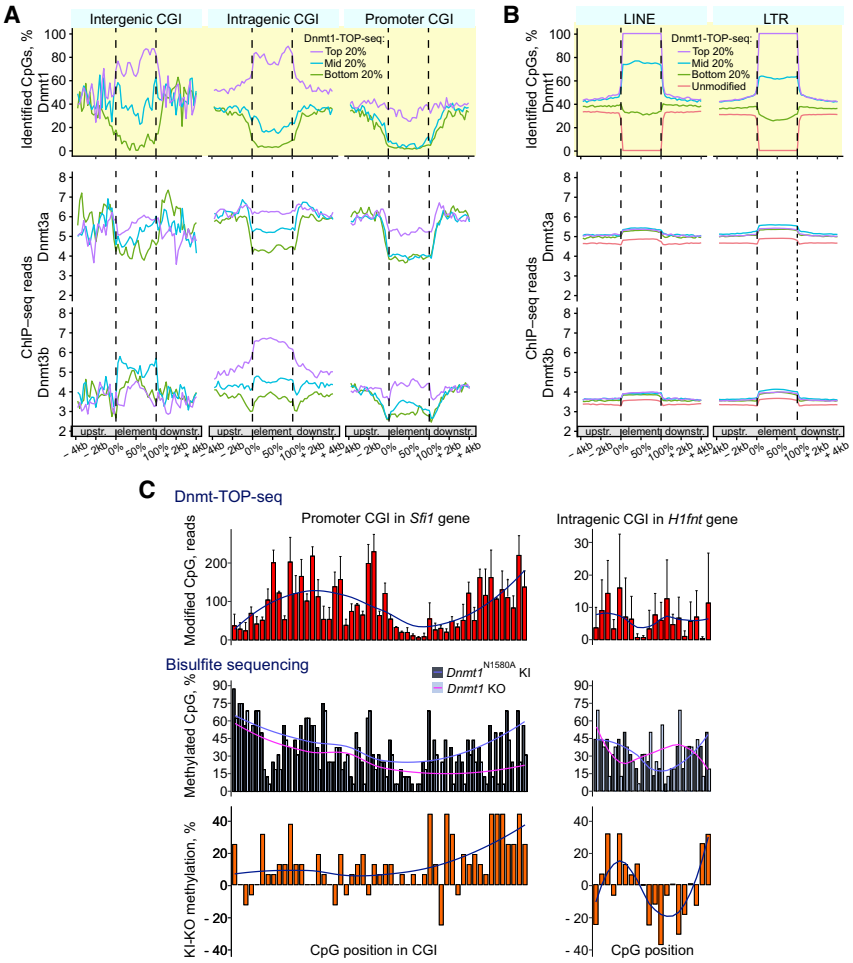


Figure 6. Contribution of Dnmt1 to methylation of genomic CpG sites in mESCs

(A and B) Comparison of Dnmt-TOP-seq CpG modification profiles (top panel in yellow) with Dnmt3a1 and Dnmt3b ChIP-seq normalized profiles (data obtained from Weinberg et al., 2019) in and around CpG islands located in promoters (2 kb upstream of protein-coding genes), intragenic and intergenic regions (A), or LINE and LTR elements (B). Profiles representing 20% slices of the most modified (top), moderately modified (mid), and least modified (bottom) regions were derived from experimental Dnmt-TOP-seq data. Bottom panel: average ChIP-seq read profiles for Dnmt3a (upper) and Dnmt3b (lower) at genomic regions selected above.

(C) Validation of CGI modification profiles in *H1fnt* and *Sfi1* genes. Columns denote read coverage or methylation levels of a particular CpG determined by Dnmt-TOP-seq (upper panel) or bisulfite sequencing (lower panel), respectively. Gaussian kernel-smoothed profiles are shown as dashed lines. Error bars denote ±SD.

the conformation of the sulfonium-bound methyl group for its in-line attack of the C5 nucleophile; although its removal leads to looser binding and positioning of AdoMet, it creates a wider exit channel for an extended transferable group out of the cofactor pocket (see Figures 1B and 7 in Lukinavičius et al., 2012). Besides obvious advantages at the genomic CRISPR-Cas engineering step (Figure S6), a single-codon replacement also offers best assurances that other biologically important functions of Dnmt1, such as allosteric regulation or interaction with other proteins, will not be affected.

Here, we also demonstrate that electropermeabilization can be used for efficient delivery of exogenous positively charged compounds into mammalian cells with no apparent size limitation. Metabolic in-cell production of AdoMet analogs has been proven for short transferrable groups (3 carbon atoms) and can generally be achieved under conditions of methionine deprivation (Hartstock et al., 2018; Shu et al., 2020). The latter may lead to dramatically altered DNA methylation, DNMT expression, and significant phenotypic changes of the stem cells (Shiraki et al., 2014; Jung et al., 2017). Our optimized pulse-labeling procedure showed negligible effects on the functionality and viability of ESCs (Figure 4). Inside the engineered cells, the internalized Ado-6-azide cofactor is utilized by the engineered Dnmt1 variant to selectively tag its biological methylation sites *in vivo*, whereas in its absence, the enzyme performs normal methylation functions using endogenous AdoMet (Figure 5). The discharge intensity, cofactor concentration, and exposure time are the major factors determining the temporal parameters and the amplitude of a labeling pulse. Our currently demonstrated density of intracellular chemical tagging was in the order of 1% of the steady-state genomic methylation level, which seemed like a reasonable trade-off between the tracking sensitivity and the preservation of a near-native state of the genome during the chase phase. The azidoalkylation density can be increased further, if necessary, by using higher concentrations of the cofactor analog (Figure 4E, left), which was currently limited by the cost and available synthetic capacity. On the other hand, the produced biallelic KI conferred a 30% reduction in global cytosine methylation (Figure S6E), which could likely be reversed by replacing only one genomic copy of *Dnmt1* in mESCs. Altogether, by further optimization of these and other parameters, the system can be tailored to meet a range of experimental demands.

The *in vivo*-tagged DNA contains CpG target sites sparsely decorated with an unnatural chemical group. Previous studies found that azidoalkylated DNA is immune to the action of restriction endonucleases *in vitro* and modification-dependent restriction systems, such as McrBC, both *in vitro* and *in vivo*; it can transform *E. coli* cells with high efficiency, indicating a good acceptance of the artificially modified DNA in bacterial cells (Lukinavičius et al., 2012, 2013). The azidoalkylated cytosines are also largely inert to TET oxidation and base excision DNA repair (Tomkuvienė et al., 2020), and they do not hinder the action of

DNA polymerases (Kriukienė et al., 2012), so they may be gradually diluted upon replication or removed via the long-patch repair mechanisms. A higher biological "durability" of the bioorthogonal chemical tags, as compared with the TET-vulnerable methyl groups, should enable direct study of the kinetic aspects of DNA methylation at single-CpG resolution, which has not been possible to date (Bachman et al., 2014; Spada et al., 2020).

The *in vivo*-generated chemical footprint of the Dnmt1 methyltransferase catalysis lends itself for precision mapping in the genomic sequence using the TOP-seq technique (Figures 5 and 6). The Dnmt-TOP-seq maps permit comprehensive high-resolution analysis of the enzyme-specific methylation landscape. This is attested by good general agreement of the inferred local and genome-wide features with BS-seq data and functional/transcription status of mESCs. On the other hand, examples of fine comparisons of the Dnmt1 catalytic footprints and the aggregate DNA methylation signal point at new intriguing facts and potential mechanistic findings. Moreover, the deposited orthogonal groups will surely be amenable to long-read DNA mapping using Oxford nanopore or single-molecule, real-time (Roberts et al., 2013; Logsdon et al., 2020) sequencing technologies, which are capable of sensing smaller epigenetic modifications on DNA. Alternatively, the deposited chemical tags could potentially be appended with fluorescent reporters for spatial nuclear mapping using super-resolution imaging (Wang et al., 2016) or used for covalent cross-linking and genomic mapping of spatially proximal sites (akin to Hi-C technology; Belton et al., 2012) for 3D tracking of the catalytic trajectories. Altogether, an inherently low selectivity of the cell electropermealization and the acceptance of a wide variety of chemical moieties including biotin by the engineered Dnmt1 (Figure 2A) provide unprecedented flexibility in chemical tracking and manipulation of epigenetic processes in live cells. Due to absolute sequence identity of the conserved catalytic motifs of the mammalian DNMT1 proteins (Figure 1C), the established approach will likely be "portable" for studies of human and other vertebrate cells. Similar engineering of other mammalian cell lines will provide powerful tools to directly track the action of individual DNMTs in a range of developmental and disease models. The new type of epigenomic information (Dnmt-selective methylation profile), which was inaccessible using previous technologies, will facilitate resolution of the many puzzles of how genomic methylation is established and maintained during development, senescence, and disease.

Limitations

As a pilot model aimed at demonstrating the general capabilities of the new approach, the described system is not optimized to suit a particular experimental study. At the DNA-tagging step, many factors such as dosage of the engineered Dnmt1, cofactor concentration, cofactor chemistry, electroporation intensity, and labeling duration can be varied to achieve desired labeling parameters and the level of deviation from a native state. To this end, although the WT and KI mESCs genetically differ only by three nucleotides, the engineered Dnmt1 is a weaker methyltransferase, leading to lower general genomic methylation levels in cells grown under LIF conditions. At the genomic mapping step, the chemical groups were shown to be well detectable and mappable using the TOP-seq technology, but other DNA sequencing modalities can also be used if single-molecule long-read analysis is desired.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Generation of murine embryonic *Dnmt1*^{N1580A} knock-in cells using CRISPR-Cas9 editing
- METHOD DETAILS
  - Methyltransferase substrates
  - Dnmt1 mutagenesis, expression and purification
  - Radiometric ^{3}H incorporation analysis
  - HPLC-MS/MS analysis of modified nucleosides
  - Single-turnover kinetic analysis
  - Comparative analysis of the Dnmt1catalytic activity with AdoMet and Ado-6-azide
  - Analysis of the Dnmt1 catalytic activity with extended AdoMet analogs
  - Two-step sequence-specific labelling of hemimethylated pΔL2-14 plasmid by copper-free click chemistry
  - Preparation of cell lysates
  - Evaluation of the Dnmt1 activity in cell lysates
  - Cell electroporation with Ado-6-azide
  - Cell viability MTT assay
  - gDNA isolation and quantitative HPLC-MS/MS
  - Preparation of Dnmt-TOP-seq libraries of mESCs
  - Sanger bisulfite sequencing
  - RT-qPCR
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Processing of Dnmt-TOP-seq data
  - Functional enrichment analysis
  - Public Sequencing Datasets

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2022.02.008.

ACKNOWLEDGMENTS

This work was supported by the European Research Council (grant ERC-AdG-2016/742654 to S.K.). The authors are grateful to A. Rukšėnaitė for assistance with MS analysis; G. Lukinavičius and E. Weinhold for kind gifts of the Ado-Buten, AdoEnYn, and AdoButyn cofactors; and E. Kriukienė for comments on the manuscript.

AUTHOR CONTRIBUTIONS

S.K. and G.V. conceived and designed the study; V.S., B.M., and L.G. performed the experiments; P.G. performed bioinformatics analysis; V.M. synthesized AdoMet analogs; V.S. and G.V. performed formal analysis; S.K. and G.V. wrote the manuscript; and all authors commented on the paper.

DECLARATION OF INTERESTS

S.K. is an inventor on patents related to mTAG labeling and TOP-seq mapping.

Received: August 13, 2021
Revised: December 4, 2021
Accepted: February 1, 2022
Published: March 3, 2022

REFERENCES

Adam, S., Anteneh, H., Hornisch, M., Wagner, V., Lu, J., Radde, N.E., Bashtrykov, P., Song, J., and Jeltsch, A. (2020). DNA sequence-dependent activity and base flipping mechanisms of DNMT1 regulate genome-wide DNA methylation. Nat. Commun. *11*, 3723.

Agrimi, G., Di Noia, M.A., Marobbio, C.M.T., Fiermonte, G., Lasorsa, F.M., and Palmieri, F. (2004). Identification of the human mitochondrial S-adenosylmethionine transporter: bacterial expression, reconstitution, functional characterization and tissue distribution. Biochem. J. *379*, 183–190.

Arand, J., Spieler, D., Karius, T., Branco, M.R., Meilinger, D., Meissner, A., Jenuwein, T., Xu, G., Leonhardt, H., Wolf, V., and Walter, J. (2012). In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases. PLoS Genet *8*, e1002750.

Bachman, M., Uribe-Lewis, S., Yang, X., Williams, M., Murrell, A., and Balasubramanian, S. (2014). 5-Hydroxymethylcytosine is a predominantly stable DNA modification. Nat. Chem. *6*, 1049–1055.

Baranauskė, S., Mickutė, M., Plotnikova, A., Finke, A., Venclovas, Č., Klimašauskas, S., and Vilkaitis, G. (2015). Functional mapping of the plant small RNA methyltransferase: HEN1 physically interacts with HYL1 and DICER-LIKE 1 proteins. Nucleic Acids Res *43*, 2802–2812.

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., and Schübeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature *520*, 243–247.

Belton, J.-M., McCord, R.P., Gibcus, J.H., Naumova, N., Zhan, Y., and Dekker, J. (2012). Hi–C: a comprehensive technique to capture the conformation of genomes. Methods *58*, 268–276.

Bestor, T.H. (2000). The DNA methyltransferases of mammals. Hum. Mol. Genet. *9*, 2395–2402.

Buitrago, D., Labrador, M., Arcon, J.P., Lema, R., Flores, O., Esteve-Codina, A., Blanc, J., Villegas, N., Bellido, D., Gut, M., et al. (2021). Impact of DNA methylation on 3D genome structure. Nat. Commun. *12*, 3243.

Dalhoff, C., Lukinavičius, G., Klimašauskas, S., and Weinhold, E. (2006). Direct transfer of extended groups from synthetic cofactors by DNA methyltransferases. Nat. Chem. Biol. *2*, 31–32.

Das, R., Lee, Y.K., Strogantsev, R., Jin, S., Lim, Y.C., Ng, P.Y., Lin, X.M., Chng, K., Yeo, G.Sh, Ferguson-Smith, A.C., and Ding, C. (2013). DNMT1 and AIM1 imprinting in human placenta revealed through a genome-wide screen for allele-specific DNA methylation. BMC Genomics *14*, 685.

Deen, J., Wang, S., Van Snick, S., Leen, V., Janssen, K., Hofkens, J., and Neely, R.K. (2018). A general strategy for direct, enzyme-catalyzed conjugation of functional compounds to DNA. Nucleic Acids Res *46*, e64.

Denis, H., Ndlovu, M.N., and Fuks, F. (2011). Regulation of mammalian DNA methyltransferases: a route to new mechanisms. EMBO Rep *12*, 647–656.

Eden, A., Gaudet, F., Waghmare, A., and Jaenisch, R. (2003). Chromosomal instability and tumors promoted by DNA hypomethylation. Science *300*, 455.

Estève, P.-O., Chang, Y., Samaranayake, M., Upadhyay, A.K., Horton, J.R., Feehery, G.R., Cheng, X., and Pradhan, S. (2011). A methylation and phosphorylation switch between an adjacent lysine and serine determines human DNMT1 stability. Nat. Struct. Mol. Biol. *18*, 42–48.

Frankish, A., Diekhans, M., Ferreira, A.-M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res *47*, D766–D773.

Gaudet, F., Hodgson, J.G., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J.W., Leonhardt, H., and Jaenisch, R. (2003). Induction of tumors in mice by genomic hypomethylation. Science *300*, 489–492.

Gerasimaitė, R., Vilkaitis, G., and Klimašauskas, S. (2009). A directed evolution design of a GCG-specific DNA hemimethylase. Nucleic Acids Res *37*, 7332–7341.

Gibas, P., Narmontė, M., Staševskij, Z., Gordevičius, J., Klimašauskas, S., and Kriukienė, E. (2020). Precise genomic mapping of 5-hydroxymethylcytosine via covalent tether-directed sequencing. PLoS Biol *18*, e3000684.

Haggerty, C., Kretzmer, H., Riemenschneider, C., Kumar, A.S., Mattei, A.L., Bailly, N., Gottfreund, J., Giesselmann, P., Weigert, R., Brändl, B., et al. (2021). Dnmt1 has de novo activity targeted to transposable elements. Nat. Struct. Mol. Biol. *28*, 594–603.

Hansen, K.D., Timp, W., Bravo, H.C., Sabunciyan, S., Langmead, B., McDonald, O.G., Wen, B., Wu, H., Liu, Y., Diep, D., et al. (2011). Increased methylation variation in epigenetic domains across cancer types. Nat. Genet. *43*, 768–775.

Hansen, R.S., Wijmenga, C., Luo, P., Stanek, A.M., Canfield, T.K., Weemaes, C.M., and Gartler, S.M. (1999). The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. Proc. Natl. Acad. Sci. USA *96*, 14412–14417.

Hartstock, K., Nilges, B.S., Ovcharenko, A., Cornelissen, N.V., Püllen, N., Lawrence-Dörner, A.-M., Leidel, S.A., and Rentmeister, A. (2018). Enzymatic or in vivo installation of propargyl groups in combination with click chemistry for the enrichment and detection of methyltransferase target sites in RNA. Angew. Chem. Int. Ed. Engl. *57*, 6342–6346.

Jeltsch, A., and Jurkowska, R.Z. (2016). Allosteric control of mammalian DNA methyltransferases - a new regulatory paradigm. Nucleic Acids Res *44*, 8556–8575.

Jung, J., Kim, L.J.Y., Wang, X., Wu, Q., Sanvoranart, T., Hubert, C.G., Prager, B.C., Wallace, L.C., Jin, X., Mack, S.C., and Rich, J.N. (2017). Nicotinamide metabolism regulates glioblastoma stem cell maintenance. JCI Insight *2*, e90019.

Jurkowska, R.Z., and Jeltsch, A. (2016). Enzymology of mammalian DNA methyltransferases. In DNA Methyltransferases - Role and Function, A. Jeltsch and R.Z. Jurkowska, eds. (Springer International Publishing)), pp. 87–122.

Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D., and Kent, W.J. (2004). The UCSC Table Browser data retrieval tool. Nucleic Acids Res *32*, D493–D496.

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science *324*, 929–930.

Kriukienė, E., Liutkevičiutė, Z., and Klimašauskas, S. (2012). 5-hydroxymethylcytosine–the elusive epigenetic mark in mammalian DNA. Chem. Soc. Rev. *41*, 6916–6930.

Lambrot, R., Xu, C., Saint-Phar, S., Chountalos, G., Cohen, T., Paquet, M., Suderman, M., Hallett, M., and Kimmins, S. (2013). Low paternal dietary folate alters the mouse sperm epigenome and is associated with negative pregnancy outcomes. Nat. Commun. *4*, 2889.

Lee, H.J., Hore, T.A., and Reik, W. (2014). Reprogramming the methylome: erasing memory and creating diversity. Cell Stem Cell *14*, 710–719.

Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J.G., Smith, A., et al. (2013). Naive pluripotency is associated with global DNA hypomethylation. Nat. Struct. Mol. Biol. *20*, 311–316.

Li, B.-Z., Huang, Z., Cui, Q.-Y., Song, X.-H., Du, L., Jeltsch, A., Chen, P., Li, G., Li, E., and Xu, G.-L. (2011). Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase. Cell Res *21*, 1172–1181.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

Liao, J., Karnik, R., Gu, H., Ziller, M.J., Clement, K., Tsankov, A.M., Akopian, V., Gifford, C.A., Donaghey, J., Galonska, C., et al. (2015). Targeted disruption

of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. Nat. Genet. *47*, 469–478.

Ling, Y., Sankpal, U.T., Robertson, A.K., McNally, J.G., Karpova, T., and Robertson, K.D. (2004). Modification of de novo DNA methyltransferase 3a (Dnmt3a) by SUMO-1 modulates its interaction with histone deacetylases (HDACs) and its capacity to repress transcription. Nucleic Acids Res *32*, 598–610.

Logsdon, G.A., Vollger, M.R., and Eichler, E.E. (2020). Long-read human genome sequencing and its applications. Nat. Rev. Genet. *21*, 597–614.

Lukinavičius, G., Lapienė, V., Staševskij, Z., Dalhoff, C., Weinhold, E., and Klimašauskas, S. (2007). Targeted labeling of DNA by methyltransferase-directed transfer of activated groups (mTAG). J. Am. Chem. Soc. *129*, 2758–2759.

Lukinavičius, G., Lapinaite, A., Urbanavičiūtė, G., Gerasimaitė, R., and Klimašauskas, S. (2012). Engineering the DNA cytosine-5 methyltransferase reaction for sequence-specific labeling of DNA. Nucleic Acids Res *40*, 11594–11602.

Lukinavičius, G., Tomkuvienė, M., Masevičius, V., and Klimašauskas, S. (2013). Enhanced chemical stability of AdoMet analogues for improved methyltransferase-directed labeling of DNA. ACS Chem. Biol. *8*, 1134–1139.

Manzo, M., Wirz, J., Ambrosi, C., Villaseñor, R., Roschitzki, B., and Baubec, T. (2017). Isoform-specific localization of DNMT3A regulates DNA methylation fidelity at bivalent CpG islands. EMBO J *36*, 3421–3434.

Martianov, I., Brancorsini, S., Catena, R., Gansmuller, A., Kotaja, N., Parvinen, M., Sassone-Corsi, P., and Davidson, I. (2005). Polar nuclear localization of H1T2, a histone H1 variant, required for spermatid elongation and DNA condensation during spermiogenesis. Proc. Natl. Acad. Sci. USA *102*, 2808–2813.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J *17*, 10–12.

Masevičius, V., Nainytė, M., and Klimašauskas, S. (2016). Synthesis of S-adenosyl-L-methionine analogs with extended transferable groups for methyltransferase-directed labeling of DNA and RNA. Curr. Protoc. Nucleic Acid Chem. *64*, 1–36, 1–1.36.13 36 31–13.

McMillan, J.M., Walle, U.K., and Walle, T. (2005). S-adenosyl-l-methionine: transcellular transport and uptake by Caco-2 cells and hepatocytes. J. Pharm. Pharmacol. *57*, 599–605.

Min, B., Park, J.S., Jeong, Y.S., Jeon, K., and Kang, Y.-K. (2020). Dnmt1 binds and represses genomic retroelements via DNA methylation in mouse early embryos. Nucleic Acids Res *48*, 8431–8444.

Ming, X., Zhang, Z., Zou, Z., Lv, C., Dong, Q., He, Q., Yi, Y., Li, Y., Wang, H., and Zhu, B. (2020). Kinetics and mechanisms of mitotic inheritance of DNA methylation and their roles in aging-associated methylome deterioration. Cell Res *30*, 980–996.

Modzelewski, A.J., Chen, S., Willis, B.J., Lloyd, K.C.K., Wood, J.A., and He, L. (2018). Efficient mouse genome engineering by CRISPR-EZ technology. Nat. Protoc. *13*, 1253–1274.

Okamoto, S., Amaishi, Y., Maki, I., Enoki, T., and Mineno, J. (2019). Highly efficient genome editing for single-base substitutions using optimized ssODNs with Cas9-RNPs. Sci. Rep. *9*, 4811.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247–257.

Osipenko, A., Plotnikova, A., Nainytė, M., Masevičius, V., Klimašauskas, S., and Vilkaitis, G. (2017). Oligonucleotide-addressed covalent 3′-terminal derivatization of small RNA strands for enrichment and visualization. Angew. Chem. Int. Ed. Engl. *56*, 6507–6510.

Peters, W., Willnow, S., Duisken, M., Kleine, H., Macherey, T., Duncan, K.E., Litchfield, D.W., Lüscher, B., and Weinhold, E. (2010). Enzymatic site-specific functionalization of protein methyltransferase substrates with alkynes for click labeling. Angew. Chem. Int. Ed. Engl. *49*, 5170–5173.

Pieters, T., and van Roy, F. (2014). Role of cell-cell adhesion complexes in embryonic stem cell biology. J. Cell Sci. *127*, 2603–2613.

Qin, W., Leonhardt, H., and Pichler, G. (2011). Regulation of DNA methyltransferase 1 by interactions and modifications. Nucleus *2*, 392–402.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308.

Roberts, R.J., Carneiro, M.O., and Schatz, M.C. (2013). The advantages of SMRT sequencing. Genome Biol *14*, 405.

Rohde, C., Zhang, Y., Reinhardt, R., and Jeltsch, A. (2010). BISMA–Fast and accurate bisulfite sequencing data analysis of individual clones from unique and repetitive sequences. BMC Bioinformatics *11*, 230.

Rols, M.-P. (2006). Electropermeabilization, a physical method for the delivery of therapeutic molecules into cells. Biochim. Biophys. Acta *1758*, 423–428.

Shiraki, N., Shiraki, Y., Tsuyama, T., Obata, F., Miura, M., Nagae, G., Aburatani, H., Kume, K., Endo, F., and Kume, S. (2014). Methionine metabolism regulates maintenance and differentiation of human pluripotent stem cells. Cell Metab *19*, 780–794.

Shu, X., Cao, J., Cheng, M., Xiang, S., Gao, M., Li, T., Ying, X., Wang, F., Yue, Y., Lu, Z., et al. (2020). A metabolic labeling method detects m6A transcriptome-wide at single base resolution. Nat. Chem. Biol. *16*, 887–895.

Sohtome, Y., and Sodeoka, M. (2018). Development of chaetocin and S-adenosylmethionine analogues as tools for studying protein methylation. Chem. Rec. *18*, 1660–1671.

Song, J., Teplova, M., Ishibe-Murakami, S., and Patel, D.J. (2012). Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. Science *335*, 709–712.

Spada, F., Schiffers, S., Kirchner, A., Zhang, Y., Arista, G., Kosmatchev, O., Korytiakova, E., Rahimoff, R., Ebert, C., and Carell, T. (2020). Active turnover of genomic methylcytosine in pluripotent cells. Nat. Chem. Biol. *16*, 1411–1419.

Staševskij, Z., Gibas, P., Gordevičius, J., Kriukienė, E., and Klimašauskas, S. (2017). Tethered oligonucleotide-primed sequencing, TOP-seq: a high-resolution economical approach for DNA epigenome profiling. Mol. Cell *65*, 554–564, e6.

Stecher, H., Tengg, M., Ueberbacher, B.J., Remler, P., Schwab, H., Griengl, H., and Gruber-Khadjawi, M. (2009). Biocatalytic Friedel-Crafts alkylation using non-natural cofactors. Angew. Chem. Int. Ed. Engl. *48*, 9546–9548.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., and Rao, A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930–935.

Tajima, S., Suetake, I., Takeshita, K., Nakagawa, A., and Kimura, H. (2016). Domain structure of the Dnmt1, Dnmt3a, and Dnmt3b DNA methyltransferases. In DNA Methyltransferases - Role and Function, A. Jeltsch and R.Z. Jurkowska, eds. (Springer International Publishing)), pp. 63–86.

Tomkuvienė, M., Ikasalaitė, D., Slyvka, A., Rukšėnaitė, A., Ravichandran, M., Jurkowski, T.P., Bochtler, M., and Klimašauskas, S. (2020). Enzymatic hydroxylation and excision of extended 5-methylcytosine analogues. J. Mol. Biol. *432*, 6157–6167.

Tomkuvienė, M., Mickutė, M., Vilkaitis, G., and Klimašauskas, S. (2019). Repurposing enzymatic transferase reactions for targeted labeling and analysis of DNA and RNA. Curr. Opin. Biotechnol. *55*, 114–123.

Troskie, R.-L., Faulkner, G.J., and Cheetham, S.W. (2021). Processed pseudogenes: a substrate for evolutionary innovation: retrotransposition contributes to genome evolution by propagating pseudogene sequences with rich regulatory potential throughout the genome. BioEssays *43*, e2100186.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S.-i., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H.R., et al. (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes Cells *11*, 805–814.

Tucker, A.M., Winkler, H.H., Driskell, L.O., and Wood, D.O. (2003). S-adenosylmethionine transport in *Rickettsia prowazekii*. J. Bacteriol. *185*, 3031–3035.

Vilkaitis, G., Suetake, I., Klimašauskas, S., and Tajima, S. (2005). Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. J. Biol. Chem. *280*, 64–72.

Wang, R., Islam, K., Liu, Y., Zheng, W., Tang, H., Lailler, N., Blum, G., Deng, H., and Luo, M. (2013). Profiling genome-wide chromatin methylation with engineered posttranslation apparatus within living cells. J. Am. Chem. Soc. *135*, 1048–1056.

Wang, S., Su, J.-H., Beliveau, B.J., Bintu, B., Moffitt, J.R., Wu, C.-t., and Zhuang, X. (2016). Spatial organization of chromatin domains and compartments in single chromosomes. Science *353*, 598–602.

Weinberg, D.N., Papillon-Cavanagh, S., Chen, H., Yue, Y., Chen, X., Rajagopalan, K.N., Horth, C., McGuire, J.T., Xu, X., Nikbakht, H., et al. (2019). The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape. Nature *573*, 281–286.

Xu, G.L., Bestor, T.H., Bourc'his, D., Hsieh, C.L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J.J., and Viegas-Péquignot, E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature *402*, 187–191.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. Omics *16*, 284–287.

Zawada, Z., Tatar, A., Mocilac, P., Buděšínský, M., and Kraus, T. (2018). Transport of nucleoside triphosphates into cells by artificial molecular transporters. Angew. Chem. Int. Ed. Engl. *57*, 9891–9895.

Zhao, L., Sun, M.-A., Li, Z., Bai, X., Yu, M., Wang, M., Liang, L., Shao, X., Arnovitz, S., Wang, Q., et al. (2014). The dynamics of DNA methylation fidelity during mouse embryonic stem cell self-renewal and differentiation. Genome Res *24*, 1296–1307.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Gelatin	Sigma-Aldrich	Cat#G1890
Trypsin-EDTA	Gibco	Cat#15400-054
Dulbecco's modified Eagle's medium	Gibco	Cat#11960-044
Embryonic stem-cell FBS	Gibco	Cat#16141-079
Penicillin-Streptomycin	Gibco	Cat#15140-122
Sodium pyruvate	Gibco	Cat#11360-070
2-mercaptoethanol	Gibco	Cat#21985-023
GlutaMAX	Gibco	Cat#35050-061
MEM Non-essential amino acids	Gibco	Cat#11140-050
Mouse leukemia inhibitory factor	EMD Millipore	Cat#ESG1106
CHIR99021	Sigma-Aldrich	Cat#SML1046
PD0325901	Sigma-Aldrich	Cat#PZ0162
Opti-MEM	Gibco	Cat#11058-021
TranscriptAid T7 High Yield Transcription Kit	Thermo Fisher Scientific	Cat#K0441
GeneArt Platinum Cas9 Nuclease	Invitrogen	Cat#B25640
Lipofectamine LTX	Invitrogen	Cat#15335-100
Puromycin	Sigma-Aldrich	Cat#P8833
Phire Tissue Direct PCR Master Mix	Thermo Fisher Scientific	Cat#F170L
S-adenosyl-L-homocysteine	Sigma-Aldrich	Cat#A9384
S-adenosyl-L-methionine	Sigma-Aldrich	Cat#A7007
AdoButen (*S*-Adenosyl-*S*-crotyl-L-homocysteine)	Prof. Elmar Weinhold (Stecher et al., 2009)	N/A
AdoButyn (*S*-Adenosyl-*S*-but-2-ynyl-L-homocysteine)	Dalhoff et al., 2006	N/A
AdoEnYn (*S*-Adenosyl-*S*-pent-4-yn-2-enyl-L-homocysteine)	Peters et al., 2010	N/A
Ado-6-azide (*S*-Adenosyl-*S*-(6-azidohex-2-ynyl)-L-homocysteine)	Masevičius et al., 2016	N/A
Ado-6-amine (*S*-Adenosyl-*S*-(6-aminohex-2-ynyl)-L-homocysteine)	Masevičius et al., 2016	N/A
Ado-6-ethyne (*S*-Adenosyl-*S*-(oct-2,6-diynyl)-L-homocysteine)	Lukinavičius et al., 2013	N/A
Ado-13-biotin	Osipenko et al., 2017	N/A
Poly(dI-dC) ● Poly(dI-dC)	Sigma-Aldrich	Cat#P4929
Poly(dG-dC) ● Poly(dG-dC)	Sigma-Aldrich	Cat#P9389
[*methyl*-^{3}H]-*S*-adenosyl-L-methionine	PerkinElmer	Cat#NET155001MC
[γ-^{32}P]ATP	PerkinElmer	Cat#NEG502Z001MC
T4 PNK	Thermo Fisher Scientific	Cat#EK0032
Geneticin	Sigma-Aldrich	Cat#A1720
PMSF	Sigma-Aldrich	Cat#78830
cOmplete Mini EDTA-free Protease Inhibitor Cocktail	Roche	Cat#4693159001
SatI restriction enzyme	Thermo Fisher Scientific	Cat#ER1641
FastDigest BspTI	Thermo Fisher Scientific	Cat#FD0834
FastDigest NotI	Thermo Fisher Scientific	Cat#FD0596
FastDigest EcoRI	Thermo Fisher Scientific	Cat#FD0274
FastDigest SacI	Thermo Fisher Scientific	Cat#FD1134

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
FastDigest HhaI	Thermo Fisher Scientific	Cat#FD1854
FastDigest XbaI	Thermo Fisher Scientific	Cat#FD0685
Scintillation cocktail Rotiszint® Eco plus	Carl Roth	Cat#0016.3
Proteinase K, recombinant, PCR grade	Thermo Fisher Scientific	Cat#EO0491
RNase A	Thermo Fisher Scientific	Cat#EN0531
Genomic DNA Clean & Concentrator-10 Kit	Zymo Research	Cat#D4011
Nuclease P1 from Penicillium citrinum	Sigma-Aldrich	Cat#N8630
FastAP Thermosensitive Alkaline Phosphatase	Thermo Fisher Scientific	Cat#EF0654
dCTP	Thermo Fisher Scientific	Cat#R0151
dGTP	Thermo Fisher Scientific	Cat#R0161
dATP	Thermo Fisher Scientific	Cat# R0141
dTTP	Thermo Fisher Scientific	Cat#R0171
5-Methyl-dCTP	Thermo Fisher Scientific	Cat#R0431
Alkyne MegaStokes dye 608	Sigma-Aldrich	Cat#79249
Igepal CA-630	Sigma-Aldrich	Cat#I8896
BCN-amine	Sigma-Aldrich	Cat#745073
MTT	Sigma-Aldrich	Cat#M5655
Fast DNA End Repair Kit	Thermo Fisher Scientific	Cat#K0771
Klenow Fragment, exo-	Thermo Fisher Scientific	Cat#EP0422
T4 DNA Ligase	Thermo Fisher Scientific	Cat#EL0011
CuBr, 99.999%	Sigma-Aldrich	Cat#254185-10G
DMSO	Sigma-Aldrich	Cat#472301
THPTA	Sigma-Aldrich	Cat#762342-500MG
dNTP Mix (2 mM each)	Thermo Fisher Scientific	Cat#R0241
Pfu DNA polymerase (recombinant)	Thermo Fisher Scientific	Cat#EP0502
Platinum SuperFi PCR Master Mix	Thermo Fisher Scientific	Cat#12358010
Dynabeads MyOne C1 Streptavidin magnetic beads	Thermo Fisher Scientific	Cat#65002
GeneJET PCR purification kit	Thermo Fisher Scientific	Cat#K0702
GeneJet NGS Cleanup kit	Thermo Fisher Scientific	Cat#K0851
DNA Clean & Concentrator-5 Kit	Zymo Research	Cat#D4013
MagJET NGS Cleanup and Size Selection Kit	Thermo Fisher Scientific	Cat# K2821
Agilent High Sensitivity DNA Kit	Agilent	Cat#5067-4626
Ion PI™ Hi-Q™ OT2 200 Kit	Thermo Fisher Scientific	Cat# A26434
Ion PI™ Hi-Q™ Sequencing 200	Thermo Fisher Scientific	Cat# A26772
Ion PI™ Chip Kit v3	Thermo Fisher Scientific	Cat# A26771
GeneJet RNA purification kit	Thermo Fisher Scientific	Cat#K0731
dsDNase	Thermo Fisher Scientific	Cat#EN0771
RevertAid™ RT reverse transcriptase	Thermo Fisher Scientific	Cat#EP0441
2x SYBR™ Green PCR Master Mix	Thermo Fisher Scientific	Cat#K253
Phusion U Hot Start polymerase	Thermo Fisher Scientific	Cat#F555S
EZ DNA Methylation-Gold Kit	Zymo Research	Cat#D5005
Deposited data		
Dnmt-TOP-seq data	This work	GEO: GSE182445
ChIP-seq data for Dnmt3a1 and Dnmt3b	Weinberg et al., 2019	GEO: GSE118785
Raw gel images	This work	Mendeley Data: 10.17632/kddxk2d7pc.1
Experimental models: Cell lines		
Mouse embryonic stem cells, E14TG2a	ATCC	CRL-1821
Mouse embryonic stem cells, E14TG2a *Dnmt1*^{N1580A}	This work	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse embryonic stem cells, E14TG2a *Dnmt1-KO*	This work	N/A
P. pastoris GS115	Thermo Fisher Scientific	Cat#C18100
Oligonucleotides		
See Table S2 for oligonucleotide sequences	Metabion	N/A
Biotinylated alkyne-containing DNA oligonucleotide, 5'-TXTTTTGTGTGGTTTGGAGACTGACTACCAGATGT AACA-Biotin; X=C8-Alkyne-dT	BaseClick	N/A
Complementary priming strand with custom LNA modifications and phosphorothioate linkages at the 3′ end, 5'-TGTTACATCTGGTAGTCAGTCTCCAAACCACACAA	Exiqon	N/A
Recombinant DNA		
pSpCas9(BB)-2A-Puro (PX459) plasmid	Addgene (Rohde et al., 2010)	Cat#62988
pΔL2-14 plasmid	Gerasimaitė et al., 2009	N/A
pUC19 plasmid	Thermo Fisher Scientific	Cat#SD0061
pPIC3.5K	Thermo Fisher Scientific	Cat#V17320
pPIC3.5K-Dnmt1-dN	Vilkaitis et al., 2005	N/A
Software and algorithms		
Cutadapt (3.4)	Martin, 2011	https://cutadapt.readthedocs.io
R (3.5)	R project	https://www.r-project.org/
FASTX (0.0.13)	Hannon Lab, CSHL	http://hannonlab.cshl.edu/fastx_ toolkit/index.html
BWA (0.7.17)	Li and Durbin, 2009	http://bio-bwa.sourceforge.net/
PyMOL (1.7)	PyMOL Molecular Graphics System by Schrödinger	https://pymol.org/
liftOver	UCSC genome browser store	https://genome-store.ucsc.edu
clusterProfile (3.10.1)	Yu et al., 2012	http://bioconductor.org/packages/ release/bioc/html/clusterProfiler.html
Other		
Repeat Masker annotation	Karolchik et al., 2004	https://genome.ucsc.edu
CpG island annotation	Karolchik et al., 2004	https://genome.ucsc.edu
Gene annotation	Frankish et al., 2019	https://www.gencodegenes.org
Detailed bench protocol	This work	Methods S1

RESOURCE AVAILABILITY

Lead contact

Further information and reasonable requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Saulius Klimašauskas (saulius.klimasauskas@bti.vu.lt).

Materials availability

Plasmids generated in this study will be made available upon request.

Mouse cell lines generated in this study will be made available upon request.

Data and code availability

- Dnmt-TOP-seq sequencing data reported in this paper have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Raw gel images have been deposited at Mendeley Data and are publicly available as of the date of publication. DOIs are listed in the key resources table.
- This paper does not report original code. All software used for analysis is listed in the key resources table and is freely available online.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mouse embryonic stem E14TG2a cell line was obtained from the American Type Culture Collection of Authenticated Cell Cultures (ATCC CRL-1821). Cells were grown on 0.15% gelatin-coated dishes in Dulbecco's modified Eagle's medium (DMEM; Gibco) containing 15% fetal bovine serum (Gibco), 1× penicillin/streptomycin (Gibco), 0.1 mM sodium pyruvate (Gibco), 0.1 mM β-mercaptoethanol (Gibco), 1 mM L-alanyl-L-glutamine (Gibco), 1× non-essential amino acids (NEAA; Gibco), 1×10^3 U/ml mouse leukemia inhibitory factor (mLIF; Millipore), 3 μM CHIR99021 (Sigma-Aldrich) and 1 μM PD0325901 (Sigma-Aldrich). Cells were maintained at 37°C in a humidified atmosphere containing 5% CO_2.

Generation of murine embryonic *Dnmt1*^{N1580A} knock-in cells using CRISPR-Cas9 editing

An E14TG2a cell line bearing a homozygous N1580A codon substitution in exon 38 of the mouse *Dnmt1* locus was generated using Cas9/sgRNA RNP nuclease and a 70-mer ssDNA as a HDR template sgRNA was designed using the Benchling tool (www.benchling.com) and synthesized using a TranscriptAid T7 High Yield Transcription Kit (Thermo Fisher Scientific) as described previously (Modzelewski et al., 2018). Before cell lipofection, Cas9/sgRNA RNP complexes were assembled in 25 μL of Opti-MEM medium (Gibco) using 500 ng of recombinant Cas9 protein (Invitrogen) and 125 ng of in vitro transcribed sgRNA at room temperature for 10 min. Subsequently, 500 ng of ssDNA template (Metabion) were added and the mixture was incubated for additional 5 min. 7.5×10^4 E14TG2a cells were transfected with the Cas9-sgRNA RNP and ssDNA mixture using Lipofectamine LTX (Invitrogen) according to the manufacturer's instructions using reverse transfection. 24 h after cell transfection, single cells were isolated into individual gelatin-coated wells in 96-well plates by serial dilution. After next 6 days of culture, single clones were picked and passaged in gelatin-coated 24-well plates. Approximately up to two next weeks during clone expansion, 1×10^5 cells of each clone were collected and screened by PCR using a Phire Tissue Direct PCR Master Mix and SatI (Thermo Fisher Scientific) restriction analysis according manufacturer's recommendations. To generate *Dnmt1* KO mESCs, E14TG2a cells were transfected with only one PX459 plasmid encoding the sgRNA sequence (500 ng) using Lipofectamine LTX as described above. Cells resistant to 2 μg/mL puromycin were plated for clonal expansion. All alterations in the *Dnmt1* coding sequence in selected clones were verified by Sanger sequencing. Primer sequences are listed in Table S2.

METHOD DETAILS

Methyltransferase substrates

AdoMet and AdoHcy were purchased from Sigma-Aldrich. [*methyl*-^{3}H]-AdoMet was acquired from PerkinElmer. AdoMet cofactor analogues were chemically synthesized from AdoHcy as previously described (Dalhoff et al., 2006; Peters et al., 2010; Lukinavičius et al., 2013; Masevičius et al., 2016; Osipenko et al., 2017) and chromatographically enriched in the enzymatically active *S*,*S*-isomer using reversed-phase preparative HPLC. For large scale *in vivo* labeling experiments, Ado-6-azide was produced as previously described (Masevičius et al., 2016) and used without reversed-phase preparative HPLC purification as a mixture or *R*,*S*- and *S*,*S*-diastereomers.

Poly(dI-dC)·poly(dI-dC) and poly(dG-dC)·poly(dG-dC) DNAs were purchased from Sigma Aldrich. Hemimethylated or non-methylated 25-mer DNA duplexes were obtained by annealing complementary oligonucleotides (Metabion). To generate radiolabeled DNA duplexes, 5′-ends of 25-mer oligonucleotides were labeled using 0.5 U/μL T4 PNK (ThermoFisher Scientific) and 0.167 μM [γ-^{32}P]ATP (PerkinElmer) for 30 min at 37°C. Duplexes were prepared by annealing 5'-labeled oligonucleotides with complementary unlabeled strands. Sequences of oligonucleotides are listed in Table S2.

pΔL2-14 plasmid DNA containing hemimethylated Gm^{5}CG/CGC sites was prepared as described previously (Gerasimaitė et al., 2009). The plasmid encodes an engineered GCG-specific DNA MTase which confers *in vivo* hemimethylation of 80 non-palindromic (GCGD, D=not C) and full (double-strand) methylation of 14 palindromic (GCGC) sites on the plasmid when expressed in bacterial cells. Briefly, *E. coli* ER2267 cells were transformed with pΔL2-14 and grown in 5 mL of LB media supplemented with 100 μg/mL ampicillin overnight at 37°C. 1/5 of cell culture was transferred in 200 mL of LB medium with 100 μg/mL ampicillin and cells were grown to OD_{600} ~0.6–0.8. 0.4 mM IPTG was added and cells incubated for 2 h at 37°C to induce the expression of the hemimethylase gene. Cells were harvested by centrifugation and plasmid DNA was purified using GeneJET Plasmid Miniprep kit (Thermo Fisher Scientific). The methylation level of the substrate DNA was estimated by digestion with m^{5}C sensitive restriction endonucleases R.HhaI (cleaves GCGC sites) and R.Bsh1236I (CGCG) followed by agarose gel electrophoresis..

Dnmt1 mutagenesis, expression and purification

In silico examination and modelling of the catalytic/cofactor binding pocket in the co-crystal structures of mouse Dnmt1-AdoHcy complexes (PDB entries: 6w8v and 6w8w) was carried out using the PyMOL Molecular Graphics System (v. 1.7).

The pPIC3.5K vector containing a His-tagged N-truncated form (residues 291–1602) of the mouse DNA methyltransferase Dnmt1 was constructed as described previously (Vilkaitis et al., 2005). Site-directed mutagenesis of the Dnmt1 variants was performed using a two-step megaprimer method as described previously (Baranauskė et al., 2015). Briefly, PCR fragments containing site-directed Dnmt1 mutations were generated during a two round PCR using primers listed in Table S2 and pPIC3.5K-Dnmt1-dN vector as a template. Obtained fragments were cloned into BsptI and NotI sites (for Q1230A substitution) or EcoRI and NotI sites (for R1576A,

N1580A and R1576A/N1580A substitution) of pPIC3.5K-Dnmt1-dN. To obtain *P. pastoris* transformants, SacI linearized plasmids were electroporated in *P. pastoris* GS115 (*his4*) strain. His^+ transformants were selected on a minimal agar medium (0.67% YNB, 2% glucose) and clones harboring a high copy number of Dnmt1 variant were subsequently selected on YPD-agar plates containing 2.5 mg/mL geneticin (Sigma).

His-tagged Dnmt1 protein variants were purified as described previously (Vilkaitis et al., 2005). Briefly, *P. pastoris* cells were grown in BMG medium (100 mM potassium phosphate, pH 6, 1.34% YNB, 0.4 ppm biotin, 1% glycerol) at 30°C until OD_{595} of 2–2.5 and then resuspended in BMM medium (100 mM potassium phosphate, pH 6, 1.34% YNB, 0.4 ppm biotin) containing 1% methanol. Following 24 h of methanol induction, cells were harvested and disrupted with a Bead Beater (Roth) in lysis buffer containing 50 mM Na_2HPO_4, 10% sucrose, 250 mM NaCl, 3 mM $MgCl_2$, 0.1% Triton X-100, 0.1 mM PMSF and cOmplete Mini EDTA-free Protease Inhibitor Cocktail (Roche), pH 7.4. Soluble fraction was collected after centrifugation for 20 min at 45000g and filtered through a 0.45 μm Millex-HV PVDF filter (Millipore). Filtered lysate was loaded onto a HiTrap IMAC HP column (Merk) preloaded with Ni^{2+} and equilibrated with wash buffer (50 mM Na_2HPO_4, 250 mM NaCl, 10% (w/v) sucrose, 3 mM $MgCl_2$, 0.1% triton X-100, 5 mM 2-mercaptoethanol, pH 6.2). Dnmt1 proteins were eluted with wash buffer supplemented with 1 M imidazole (pH 6.2) and dialyzed twice against buffer A overnight and against storage buffer (50 mM Na_2HPO_4, 250 mM NaCl, 0.1% triton X-100, 5 mM 2-mercaptoethanol and 50% glycerol, pH 7.4). Aliquoted proteins were stored at -20°C

Radiometric ^3H incorporation analysis

Dnmt1 catalytic parameters were determined by measuring the incorporation of ^3H-methyl group using [*methyl*-^3H]-AdoMet as described previously (Vilkaitis et al., 2005). Briefly, reactions were performed in reaction buffer (20 mM Tris-HCl, pH 7.4, 0.5 mM EDTA, 0.2 mM DTT, 0.1 mg/mL bovine serum albumin) containing 20 nM WT or 100 nM engineered Dnmt1, 5.3 μM [*methyl*-^3H]-AdoMet, 6 μM DNA substrate (poly(dI-dC)·(dI-dC) or poly(dG-dC)·(dG-dC), hemimethylated or non-methylated 25-mer duplex). All reactions were conducted for 40 min at 37°C and quenched by heating for 10 min at 80°C. In separate series substrate concentrations were varied in the range of 0.02–11 μM for cofactor, or 0.002–12 μM (target sites) for DNA. To determine the levels of ^3H-methylcytosine, samples were spread on 2.3 cm DE-81 filter papers (Whatman, UK), washed four times with 50 mM $NaPO_4$ buffer (pH 7.4), twice with H_2O, twice with 96% ethanol, once with acetone and dried under IR lamp. ^3H-methyl group incorporation was measured in 2 mL of scintillation cocktail Rotiszint® Eco plus (Carl Roth) using a Hidex 300SL (LabLogic Systems) scintillation spectrometer.

HPLC-MS/MS analysis of modified nucleosides

Typically, 50 ng of modified DNA was digested to nucleotides with 0.01 U/μL of nuclease P1 (Sigma) in 40 μL of P1 reaction buffer (10 mM NaAc, 1 mM ZnAc, pH 5.5) for 4 h at 50°C followed by an overnight incubation with 0.01 U/μL of FastAP alkaline phosphatase (Thermo Fisher Scientific) at 37°C. Reactions were quenched by heating at 80°C for 10 min and centrifugation at 19000*g* for 30 min at 4°C. Centrifugate samples were loaded on an integrated HPLC/ESI-MS/MS system (Agilent 1290 Infinity/ 6410 Triple Quad LC/MS) equipped with a Supelco Discovery HS C18 column (7.5 cm × 2.1 mm, 3 μm) by elution with a linear gradient of solvents A (0.0075% formic acid in water) and B (0.0075% formic acid in acetonitrile) at a flow of 0.3 mL/min at 30°C as follows: 0–5 min, 0% B; 5–15 min, 10% B; 15–20 min, 100% B. Mass spectrometer was operating in the positive ion MRM mode and intensities of nucleoside-specific ion transitions were recorded: N_3-m^5dC 349.2→233.1, m^5dC 242.1→126.1; dC 228.1→112.0, dG 268.1→152.1.

For quantitative analysis, standard curves were generated using FastAP-treated nucleotides (Thermo Fisher Scientific): m^5dC (linear range of 0.0125–0.8 pmol) and dC, dG, dT, dA (0.38–12 pmol). N_3-m^5dC calibration was achieved by analysis of 0.2–25 fmol of hydrolysates obtained from hemimethylated 25-mer oligonucleotide duplex that was fully modified with Dnmt1 N1580A protein in the presence of Ado-6-azide, and thus contained equimolar amounts of m^5dC and N_3-m^5dC). To compensate for DNA input variations, all signals were normalized to dG. Data were analyzed using Agilent MassHunter software and Microsoft Excel.

Single-turnover kinetic analysis

Single turnover reactions containing 100 nM Dnmt1, 50 nM appropriately 5'-^{32}P labeled 25-mer oligonucleotide duplex and 100 μM cofactor (AdoMet or Ado-6-azide) in reaction buffer (20 mM Tris-HCl, pH 7.4, 0.5 mM EDTA, 0.2 mM DTT, 0.1 mg/mL bovine serum albumin) were incubated for specified periods of time and quenched by heating for 10 min at 80°C. Control reactions contained no Dnmt1. Resulting modified DNA was reannealed with a 125-molar excess of an unmodified complementary 25-mer strand and digested with the R.HhaI endonuclease (ThermoFisher Scientific) in FastDigest buffer for 15 min at 37°C. Samples were fractionated on a 13% denaturing polyacrylamide gel with 7 M urea under denaturing conditions. Gels were exposed to Phosphor imaging plates and scanned by FLA-5100 Image Reader (Fujifilm) with a red 635 nm laser and IP filter and analyzed using Multi Gauge software (Fujifilm).

Comparative analysis of the Dnmt1catalytic activity with AdoMet and Ado-6-azide

Catalytic reactions were carried out with 100 nM Dnmt1 (WT, N1580A and R1576A/N1580A) 50 nM 5'-^{32}P labeled hemimethylated duplex, and 100 μM cofactor (AdoMet and/or Ado-6-azide premixed at certain ratios) for 1 h at 37°C and then enzyme was inactivated by heating for 10 min at 80°C. To differentiate the migration of methylated and azidoalkylated strands in polyacrylamide gels, "click" modification of the latter strand was performed by supplementing reactions with 1 mM BCN-amine (Sigma) for 3 h at 37°C. Samples were next proceeded as described above.

Analysis of the Dnmt1 catalytic activity with extended AdoMet analogs

DNA modification reactions were performed using 50 nM 5'-^{32}P labeled hemimethylated 25-mer oligonucleotide duplex, 100 nM Dnmt1 variant in the presence of 100 μM AdoMet or synthetic cofactors in reaction buffer (20 mM Tris-HCl, pH 7.4, 0.5 mM EDTA, 0.2 mM DTT, 0.1 mg/mL bovine serum albumin) for 1 h at 37°C quenched by heating for 10 min at 80°C. Samples were processed as described above.

Two-step sequence-specific labelling of hemimethylated pΔL2-14 plasmid by copper-free click chemistry

500 ng of pΔL2-14 plasmid DNA was supplemented with 100 nM Dnmt1 and 100 μM Ado-6-azide in reaction buffer (20 mM Tris-HCl, pH 7.4, 0.5 mM EDTA, 0.2 mM DTT, 0.1 mg/mL bovine serum albumin) and incubated for 1 h at 37°C. Reactions were stopped by heating for 10 min at 80°C, supplemented with 1% SDS, treated with 1 μg/mL of Proteinase K for 30 min at 55°C and purified by ethanol precipitation. Modified plasmid DNA was then treated with 5 μM alkyne MegaStokes 608 (Sigma) in 20 mM of Tris-HCl buffer (pH 7.4) for 3 h at 37°C and precipitated with ethanol. 300 ng of labelled DNA was fragmented with EcoRI and XbaI (ThermoFisher Scientific) and analyzed on a 1% agarose gel. Gels were scanned for MegaStokes 608 fluorescence in a FLA-5100 Image Reader using a 473 nm laser. To visualize bulk DNA, gels were stained with EtBr, rescanned and analyzed using Multi Gauge software.

Preparation of cell lysates

10^6 murine embryonic stem cells were collected and washed once in 1 ml of PBS, once in 500 μL wash buffer (20 mM Tris-HCl, pH 7.4, 20 mM NaCl, 6 mM $MgCl_2$) and spun down for 2 min at 1000 g at 4°C. Cell pellet was gently suspended by pipetting in 500 μL hypotonic buffer (20 mM Tris-HCl, pH 7.4, 20 mM NaCl, 6 mM $MgCl_2$, 1 mM DTT, 0.1% Igepal CA-630). After incubation for 10 min in ice, cells were spun down for 5 min at 1000 g at 4°C. Cell pellet was gently resuspended in 100 μL of nuclear extraction buffer (60 mM Tris-HCl, pH 7.4, 1.2 mM EDTA, 0.5% Igepal CA-630, 1 mM DTT, 0.1 mg/mL bovine serum albumin, cOmplete Mini EDTA-free Protease Inhibitor Cocktail) and sonicated 5 times (30 s pulse and 30 s pause) at 4°C in a Bioruptor sonicator (Diagenode). Soluble cell lysates were collected following centrifugation for 15 min at 14000 g at 4°C.

Evaluation of the Dnmt1 activity in cell lysates

In vitro Dnmt1 methylation activity in freshly prepared cell lysates was determined by measuring incorporation of ^{3}H-methyl group from [*methyl*-^{3}H]-AdoMet into DNA. Methylation reactions were performed in 20 μL of E14TG2a cell lysate containing 100 nM Dnmt1, 5.3 μM [*methyl*-^{3}H]-AdoMet and 6 μM DNA. In vitro alkylation reactions were performed in 20 μL of E14TG2a cell lysate containing 100 nM Dnmt1, 100 μM Ado-6-azide and 500 ng of pΔL2-14 plasmid DNA. For the evaluation of endogenous Dnmt1 activity, alkylation reactions were performed as above with no exogenous Dnmt1 protein added. All reactions were conducted for 1 h at 37°C and stopped by heating for 10 min at 80°C. Alkylation reactions were then supplemented with 1% SDS and treated with 1 μg/mL Proteinase K for 30 min at 55°C and modified DNA was purified by phenol-chloroform extraction.

Cell electroporation with Ado-6-azide

Cells were plated in 24-well plates at a density of $5x10^4$ cells per well and incubated for 48 h in DMEM media supplemented with fetal bovine serum and mLIF. Each well was washed twice with Opti-MEM medium (ThermoFisher Scientific) and inoculated with freshly prepared 300 μL of 0.5–1 mM Ado-6-azide solution in Opti-MEM medium and electroporated using a NEPA21 electroporator (Nepa Gene) fitted with a cell culture plate electrode CUY900 at a cell poring pulse voltage of 200 V and a pulse duration of 5 ms. After electroporation, cells were incubated for further 25 min at 37°C, washed with Opti-MEM, covered with fresh mESC medium and incubated for further 1–6 hours. Cell pellets from each sample were collected and stored at -20°C.

Cell viability MTT assay

Cells were treated with 500 μL of 0.1 mg/mL 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT, Sigma) in PBS for 3 min at 37°C, washed twice with PBS and lysed with 200 μL neat isopropanol. Formazan absorption in samples was measured at 570 nm using a Synergy H4 (Biotek) plate reader.

gDNA isolation and quantitative HPLC-MS/MS

$5x10^5$ of pelleted embryonic stem cells were resuspended in 200 μL of a 1:1 mixture of PBS and Lysis buffer (0.2 M Tris-HCl, 0.25 M NaCl, 0.025 M EDTA, 0.5% SDS, pH 8; 1 mg/mL Proteinase K) and incubated for 40 min at 55°C. Samples were then treated with RNase A (ThermoFisher Scientific) for 10 min at room temperature and genomic DNA isolated using a Genomic DNA Clean & Concentrator-10 Kit (Zymo Research) and stored at -20°C. 500 ng of isolated gDNA was treated with P1 endonuclease and FastAP phosphatase and subjected to MS/MS-HPLC spectrometer as described above.

Preparation of Dnmt-TOP-seq libraries of mESCs

The preparation of Top-seq libraries was carried out as described (Staševskij et al., 2017). Briefly, 900 ng of genomic DNA, isolated from WT or *Dnmt1*^{N1580A} E14TG2a cells electroporated with 1 mM Ado-6-azide, was sheared to 200 bp with a Covaris E220

sonicator. Fragmented DNA was end-repaired using a DNA End Repair Kit (Thermo Fisher Scientific) as recommended by the manufacturer and purified using a GeneJet PCR Purification Kit (Thermo Fisher Scientific). Next, DNA fragments were 3'-A-tailed with Klenow exo- polymerase in Klenow buffer (ThermoFisher Scientific) in the presence of 0.5 mM dATP at 37°C for 45 min followed by inactivation at 75°C for 15 min and purified through a DNA Clean & Concentrator-5 column (Zymo Research). Annealed partially complementary A1/A2 adapters (Metabion) were ligated using 15 U of T4 DNA Ligase (ThermoFisher Scientific) in ligase buffer at 22°C overnight following by thermal inactivation at 65°C for 10 min and purification through a DNA Clean & Concentrator-5 column (Zymo Research). Next, DNA was supplemented with 20 μM biotinylated alkyne-containing DNA oligonucleotide (5′-T(alkyneT) TTTTGTGTGGTTTGGAGACTGACTACCAGATGTAACA-(biotin)-3′; Base-click) and 8 mM CuBr: 24 mM THPTA mixture (Sigma) in 50% of DMSO, incubated for 20 min at 45°C, and diluted to ~1% DMSO before purification by a GeneJet NGS Cleanup kit (ThermoFisher Scientific). Biotinylated DNA was enriched by incubation with 0.1 mg Dynabeads MyOne C1 Streptavidin (ThermoFisher Scientific) in buffer A (10 mM Tris–HCl (pH 8.5), 1 M NaCl) at room temperature for 3 h on a roller. After incubation, DNA-bound beads were washed twice with buffer B (10 mM Tris–HCl (pH 8.5), 3 M NaCl, 0.05% Tween 20); twice with buffer A (supplemented with 0.05% Tween 20); once with 100 mM NaCl, resuspended in water and heated for 5 min at 95°C to elute the enriched DNA fraction. Recovered DNA was subsequently applied to a priming reaction containing 1 U Pfu DNA polymerase (ThermoFisher Scientific), 0.2 mM dNTP and 0.5 μM complementary priming strand 5′-TGTTACATCTGGTAGTCAGTCTCCAAACCACACAA-3′ (with custom LNA modifications and phosphorothioate linkages at the 3′ end; Exiqon). The priming reaction was performed at the following cycling conditions: 95°C 2 min; 5 cycles at 95°C 1 min, 65°C 10 min, 72°C 10 min. Amplification of a primed DNA library was carried out by adding the priming reaction mixture to 50 μL of amplification reaction containing 25 μL of 2 × Platinum SuperFi PCR Master Mix (ThermoFisher Scientific) and barcoded fusion PCR primers A(Ad)-EP-barcode-primer (63 nt) and trP1(Ad)-A2-primer (45 nt) at 0.5 μM each (both primers contained phosphorothioate modifications). Thermocycler conditions were as follows: 94°C 4 min; 15 cycles at 95°C 1 min, 60°C 1 min, 72°C 1 min. Final libraries were size-selected for ~ 300 bp fragments using a MagJET NGS Cleanup and Size Selection Kit (ThermoFisher Scientific). Library quality and quantity were assessed with a 2100 Bioanalyzer (Agilent). Libraries were subjected to Ion Proton (ThermoFisher Scientific) sequencing. Primer sequences are enlisted in Table S2.

Sanger bisulfite sequencing

400 ng of genomic DNA isolated from ES cells was bisulfite-converted using a Zymo DNA Methylation Gold Kit (Zymo Research) and then subjected to PCR to amplify target regions using Phusion U Hot Start polymerase (Thermo Fisher Scientific). Both procedures were performed according to manufacturer's instructions. Obtained PCR products were purified using a GeneJET PCR purification kit (Thermo Fisher Scientific) and cloned into the pUC19 vector. Sixteen clones of each sample were randomly selected for Sanger sequencing. CpG methylation status of bisulfite sequencing data was evaluated using BISMA tool (Rohde et al., 2010). All primers are listed in Table S2.

RT-qPCR

$5x10^5$ ES cells were processed using a GeneJet RNA purification kit (Thermo Fisher Scientific) and then treated with dsDNAse (Thermo Fisher Scientific). cDNA synthesis was performed with 500 ng of total RNA and RevertAid RT reverse transcriptase (Thermo Fisher Scientific). Quantitative real-time PCR analysis was carried out on a Rotor-Gene 6000 (Corbett Life Science) using a SYBR™ Green PCR Master Mix (Thermo Fisher Scientific). All protocols were carried out according to manufacturer's recommendations. Determined mRNA levels were normalized to *Gapdh* expression. PCR primers (Metabion) are listed in Table S2.

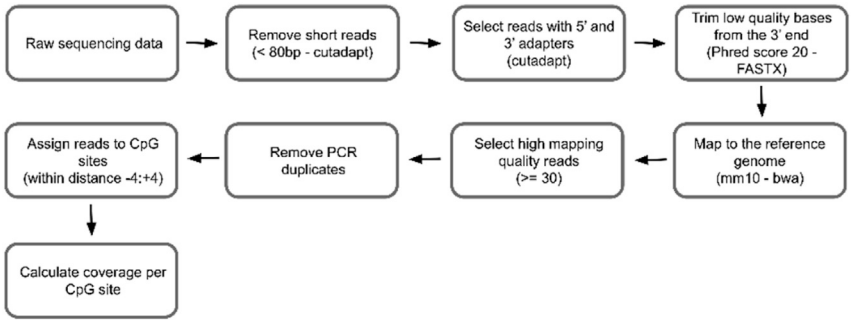
QUANTIFICATION AND STATISTICAL ANALYSIS

The biochemical data analysis and statistical tests used are described within the figure legends and methods section describing each experiment.

Processing of Dnmt-TOP-seq data

mESC TOP-seq data was processed as described previously (Gibas et al., 2020). First, reads longer than 80 nt (cutadapt -1 80) and containing 5' adapter sequences were retained (cutadapt -g ^GTTACATCTGGTAGTCAGTCTCCAAACCACACAAAAA -e 0.1 -O 10), followed by trimming of 3' adapter sequence using cutadapt tool (cutadapt -a AGATTGGAAGAGTGTCATGTAGGGAAAGAGTG -e 0.1 -O 10) (Martin, 2011). FASTX quality trimmer was then used to trim 3' read ends below a phred quality score of 20 (fastq_quality_trimmer -t 20 -l 15) (http://hannonlab.cshl.edu/fastx_toolkit/index.html). Processed read sequences were mapped to the mouse genome (build mm10) using BWA tool and only mappings with quality score 30 or above were retained (Li and Durbin, 2009). Identical reads (the same original length and starting genomic coordinate) were termed PCR duplicates and only one per group was retained. All selected reads were assigned to a particular CpG site with absolute distance between the read start and CpG site between 0 and 4 bp using an in-house script. Reads with identical distance to multiple CpG sites were omitted from analysis. Finally, CpG coverage was defined as a sum of all assigned reads on both strands.

Data analysis pipeline is provided in the following scheme:



Functional enrichment analysis

Modified promoter CpG island Gene Ontology (GO) analysis. All CGIs intersecting with a protein-coding gene upstream region (2 kb) were defined as promoter CGIs. For each biological replicate fraction of CpG sites with at least one read was computed and only CGIs where all biological replicates had identified CpG sites were selected. Next, GO Enrichment Analysis of selected genes was performed using the enrichGO function from the clusterProfile suite (Yu et al., 2012). All protein-coding genes were used as a background set and only GO enrichments with Benjamini-Hochberg adjusted q-value less than 0.001 were selected.

Enrichment of various genomic elements with the modification signal was computed using Fisher's exact test. First, a contingency table was formatted for each CpG site falling within a specific genomic element and containing at least one read. Next, Fisher's exact test was performed to estimate the odds ratio and significance for each enrichment or depletion. Modification enrichment within protein-coding gene bodies was calculated in a similar fashion. For each gene, a contingency table was computed to test if a CpG site contains at least one read and overlaps a specific gene. Fisher's exact test was used to calculate the odds ratio and p-value. Next, genes that showed the same direction of change (either depletion or enrichment) and FDR adjusted q-value less than 0.05 in all biological replicates were defined as depleted or enriched with DNA modification. Finally, GO enrichment analysis was performed as mentioned above and only genes with Benjamini-Hochberg adjusted q-value less than 0.05 were selected.

Mouse genome sequence in the fasta file format, Repeat element (Repeat Masker annotation, rmsk) and CpG island (cpgIslandExt annotation) genomic locations in the BED file format were downloaded from the UCSC database (Karolchik et al., 2004). Reference gene annotations was obtained from the GENCODE project in the GTF file format (Frankish et al., 2019).

Public Sequencing Datasets

Dnmt-TOP-seq data have been deposited at the NCBI GEO under accession number GSE182445. ChIP-seq data for mouse Dnmt3a1 and Dnmt3b were downloaded through NCBI GEO GSE118785 (Weinberg et al., 2019).