nature biotechnology

SUPPLEMENTARY INFORMATION

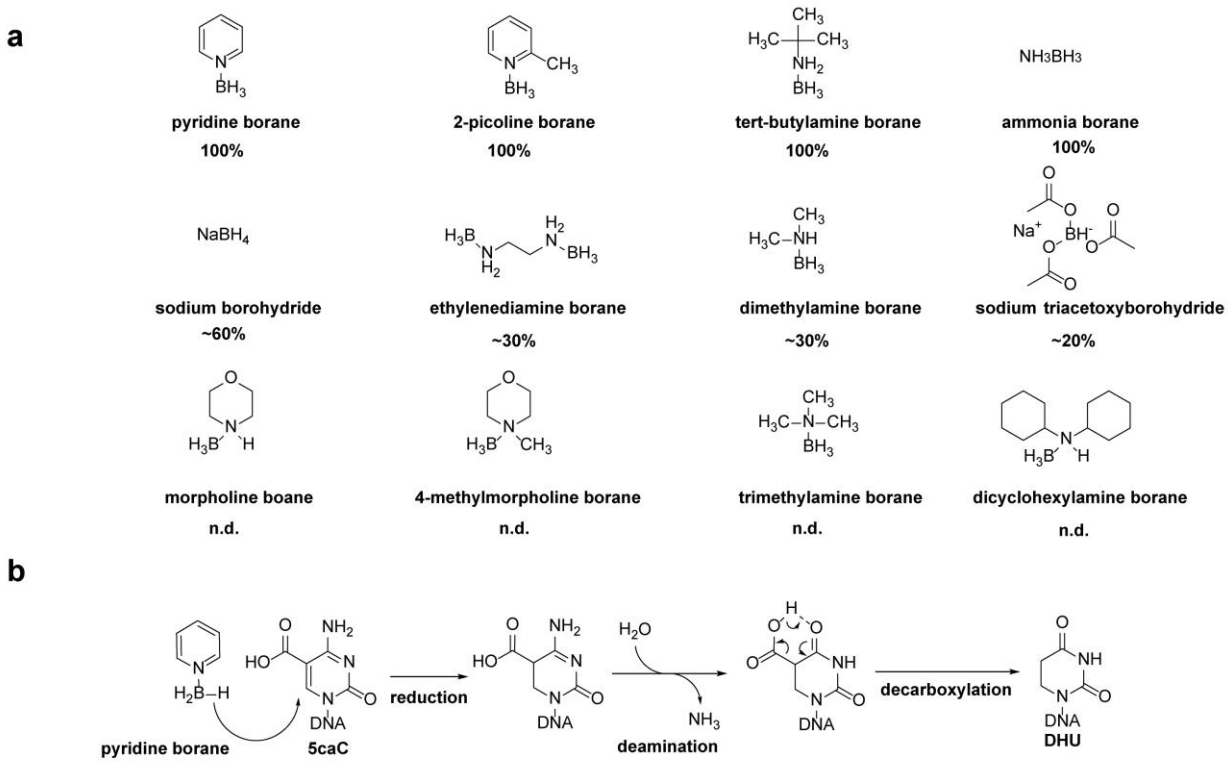
https://doi.org/10.1038/s41587-019-0041-2

In the format provided by the authors and unedited.

Bisulfite-free direct detection of 5-methylcytosine and 5-hydroxymethylcytosine at base resolution

Yibin Liu[1,2,6], Paulina Siejka-Zielińska[1,2,6], Gergana Velikova[1,2], Ying Bi[1,2], Fang Yuan[1,2,3], Marketa Tomkova[1], Chunsen Bai[1,2,4], Lei Chen[1,2,5], Benjamin Schuster-Böckler[1]* and Chun-Xiao Song[1,2]*

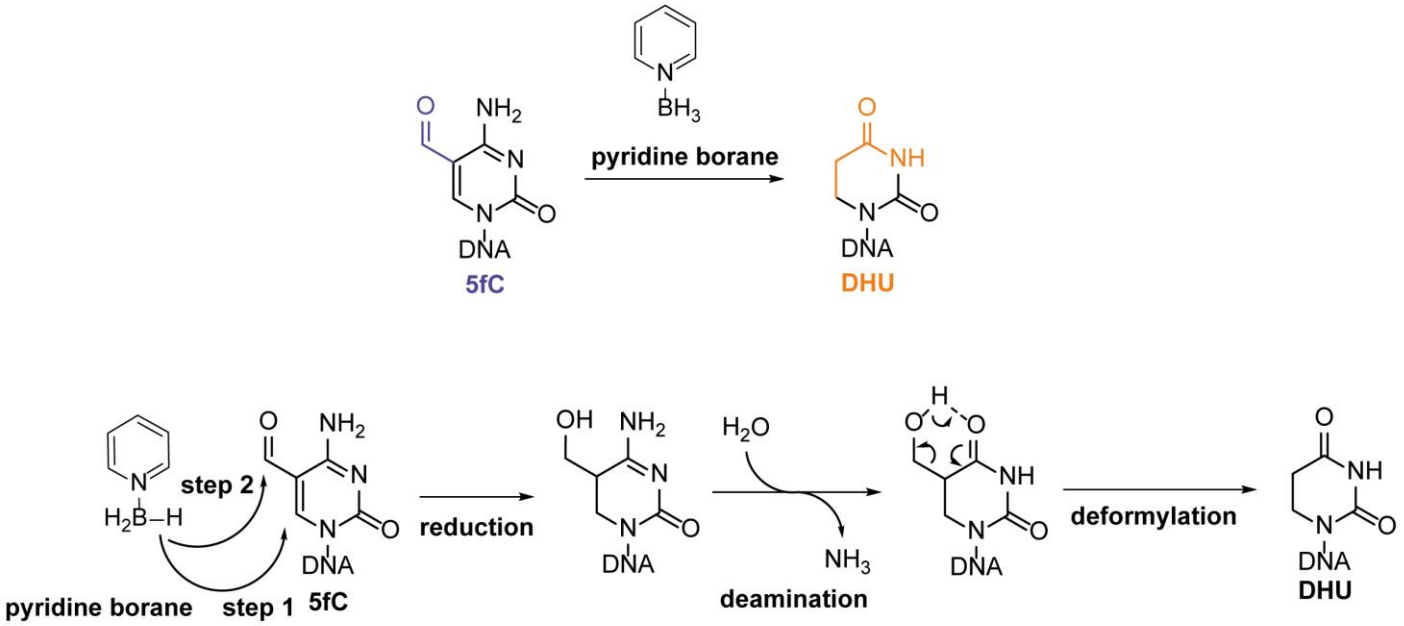
[1]Ludwig Institute for Cancer Research, Nuffield Department of Medicine, University of Oxford, Oxford, UK. [2]Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK. [3]Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, Peking University, Beijing, China. [4]College of Chemistry, Nankai University, Tianjin, China. [5]Center for Mitochondrial Biology and Medicine and Center for Translational Medicine, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, China. [6]These authors contributed equally: Yibin Liu, Paulina Siejka-Zielińska *e-mail: benjamin.schuster-boeckler@ludwig.ox.ac.uk; chunxiao.song@ludwig.ox.ac.uk



Supplementary Figure 1

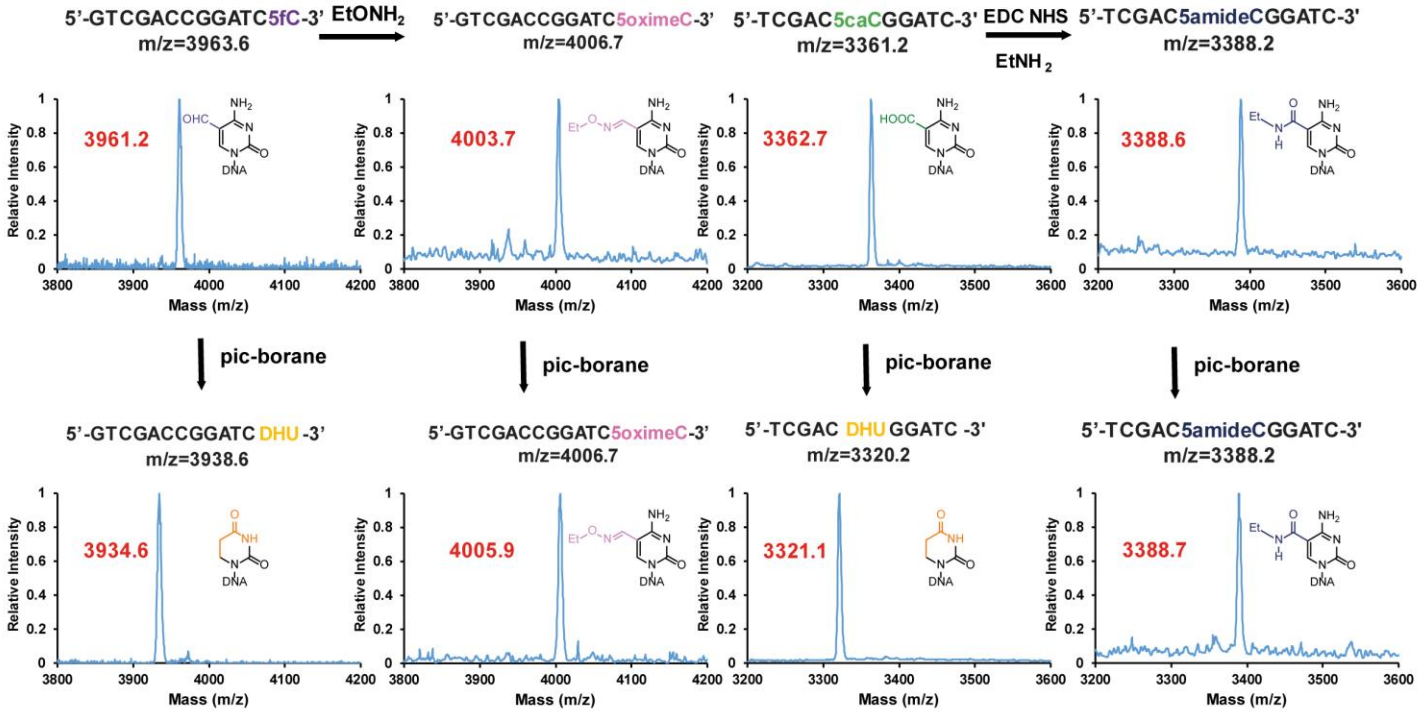
Borane-containing compounds screening and proposed mechanism for the borane reaction of 5caC.

(**a**) Borane-containing compounds screened for conversion of 5caC to DHU in an 11mer oligo, with conversion rate estimated by MALDI. 2-picoline borane (pic-borane), borane pyridine, tert-butylamine borane, and ammonia borane could completely convert 5caC to DHU while ethylenediamine borane and dimethylamine borane only gave around 30% conversion rate. No detectable products were measured (n.d.) with morpholine borane, 4-methylmorpholine borane, trimethylamine borane, and cyclohexylamine borane. Other reducing agents such as sodium borohydride and sodium tri(acetoxy)borohydride decomposed rapidly in acidic media and led to incomplete conversion. Sodium cyanoborohydride was not used due to potential for hydrogen cyanide formation under acidic conditions. Pic-borane and pyridine borane were chosen because of the complete conversion, low toxicity and high stability. (**b**) Proposed mechanism for the borane reaction of 5caC to DHU.



Supplementary Figure 2

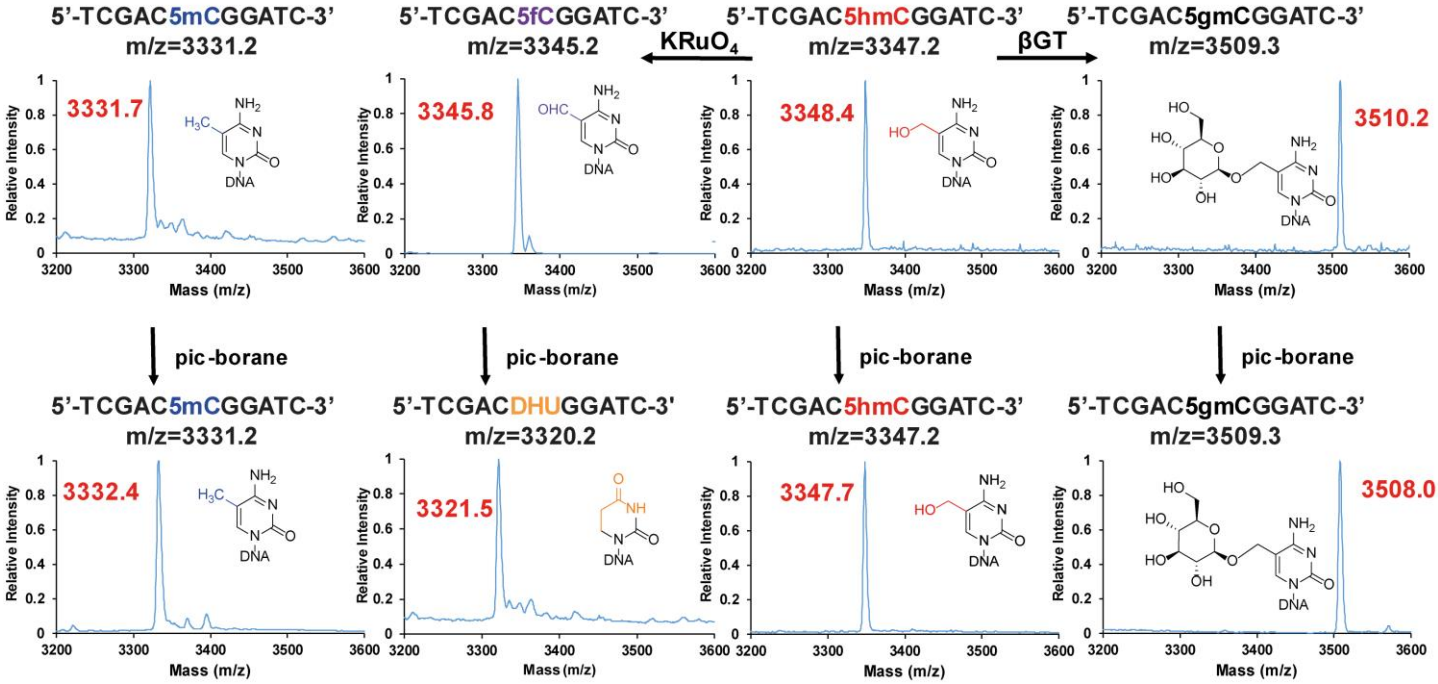
Proposed mechanism for the borane reaction of 5fC to DHU.



Supplementary Figure 3

MALDI characterization of 5fC and 5caC containing model DNA oligos treated by pic-borane with or without the blocking of 5fC and 5caC.

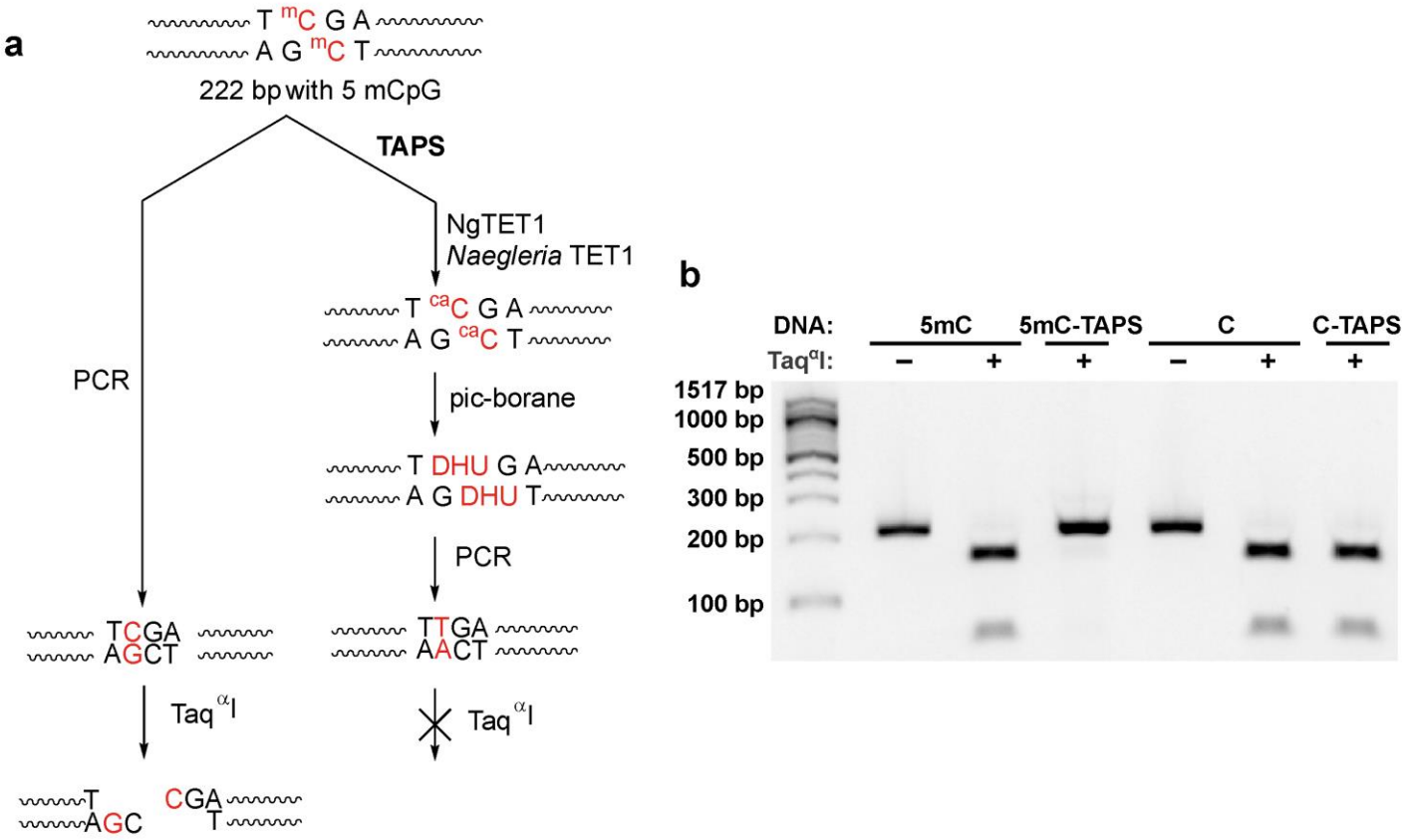
5fC was blocked by O-ethylhydroxylamine which becomes oxime and resists pic-borane conversion while 5caC was blocked by ethylamine via EDC conjugation and converted to amide which blocks conversion by pic-borane. All experiments were performed once. Calculated MS was shown in black, and observed MS was shown in red.



Supplementary Figure 4

MALDI characterization of 5mC and 5hmC containing model DNA oligos treated by $KRuO_4$ and pic-borane with or without blocking of 5hmC.

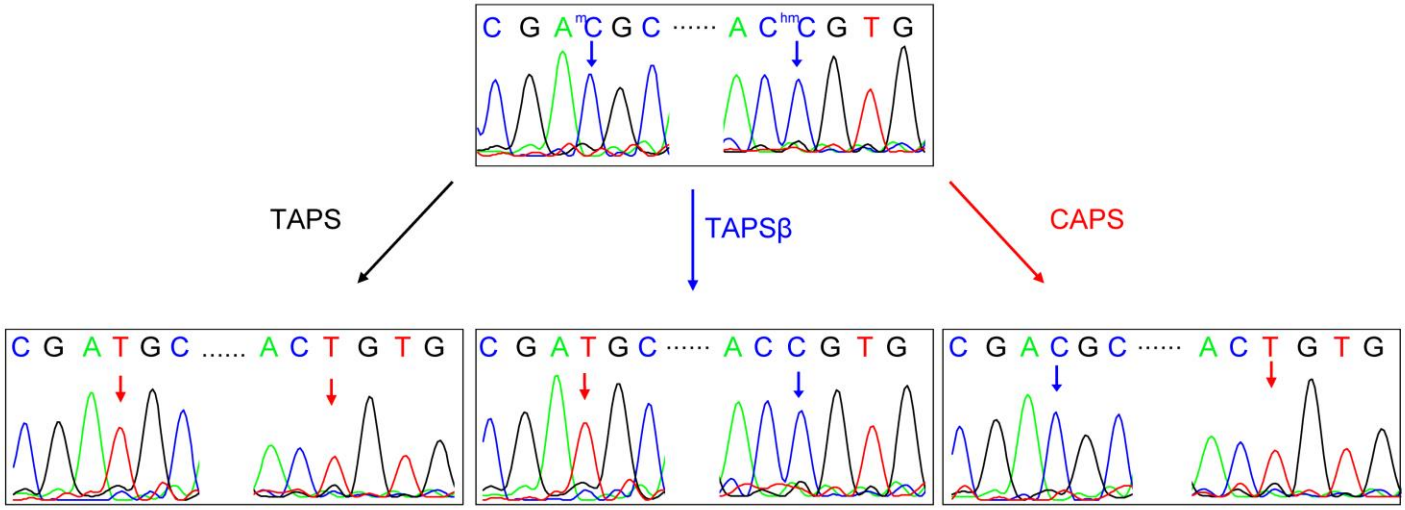
5hmC could be blocked by βGT with glucose and converted to 5gmC. 5mC, 5hmC and 5gmC could not be converted by pic-borane. 5hmC could be oxidized by $KRuO_4$ to 5fC, and then converted to DHU by pic-borane. All experiments were performed once. Calculated MS was shown in black, observed MS was shown in Red.



Supplementary Figure 5

Restriction enzyme digestion showed TAPS effectively converts 5mC to T.

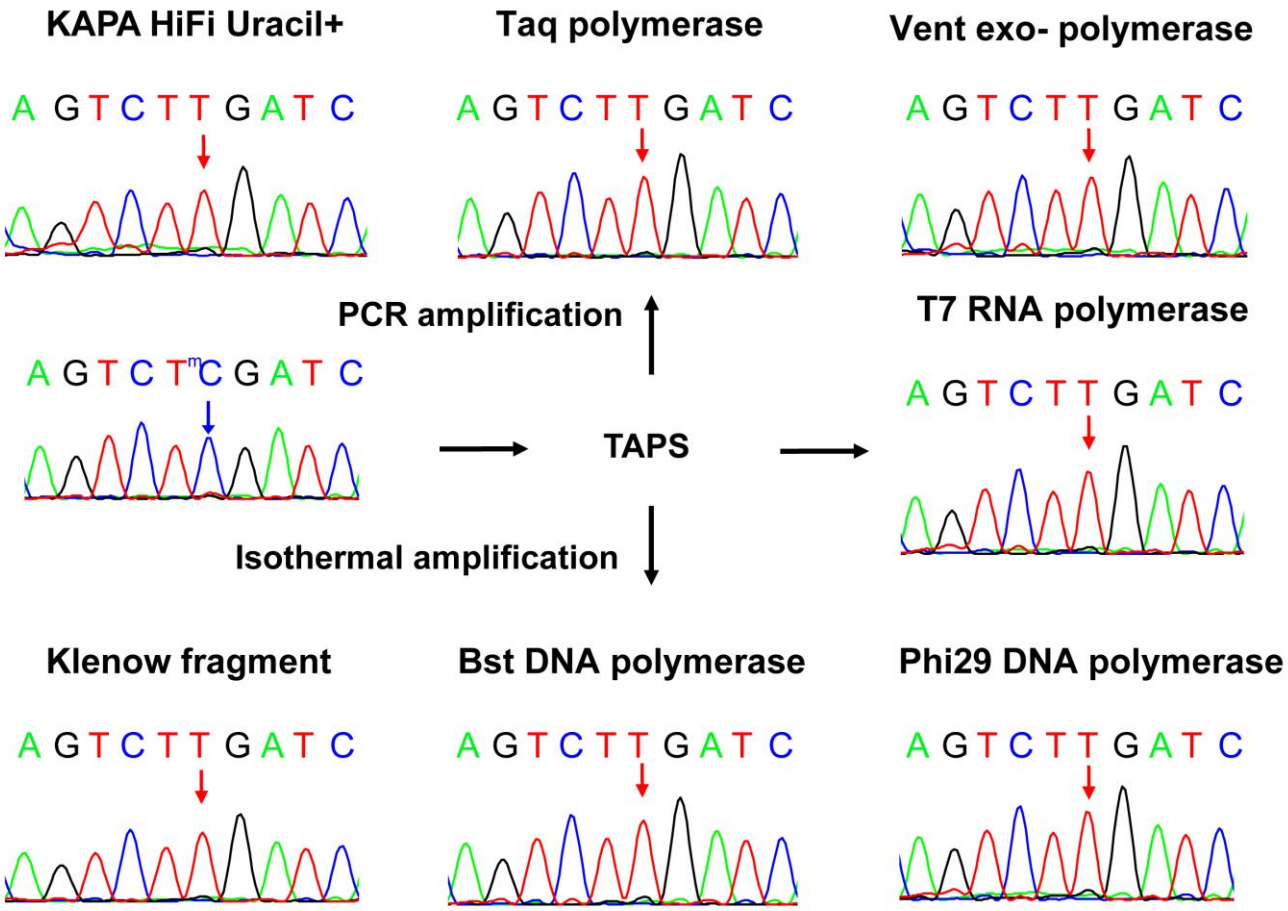
(**a**) Illustration of restriction enzyme digestion assay to confirm the sequence change caused by TAPS. (**b**) Taq$^{\alpha}$I-digestion tests to confirm the C-to-T transition caused by TAPS. A PCR-amplified 222 bp model DNA with Taq$^{\alpha}$I restriction site in the middle can be cleaved, whereas the amplified product of 5mC-TAPS stayed intact, suggesting loss of the restriction site and hence C-to-T transition. TAPS did not result in C-to-T transition on the unmethylated cytosine since C-TAPS was cleaved in the same way as the original untreated C. Experiment was performed once.



Supplementary Figure 6

Complete C-to-T transition induced after TAPS, TAPSβ and CAPS as indicated by Sanger sequencing.

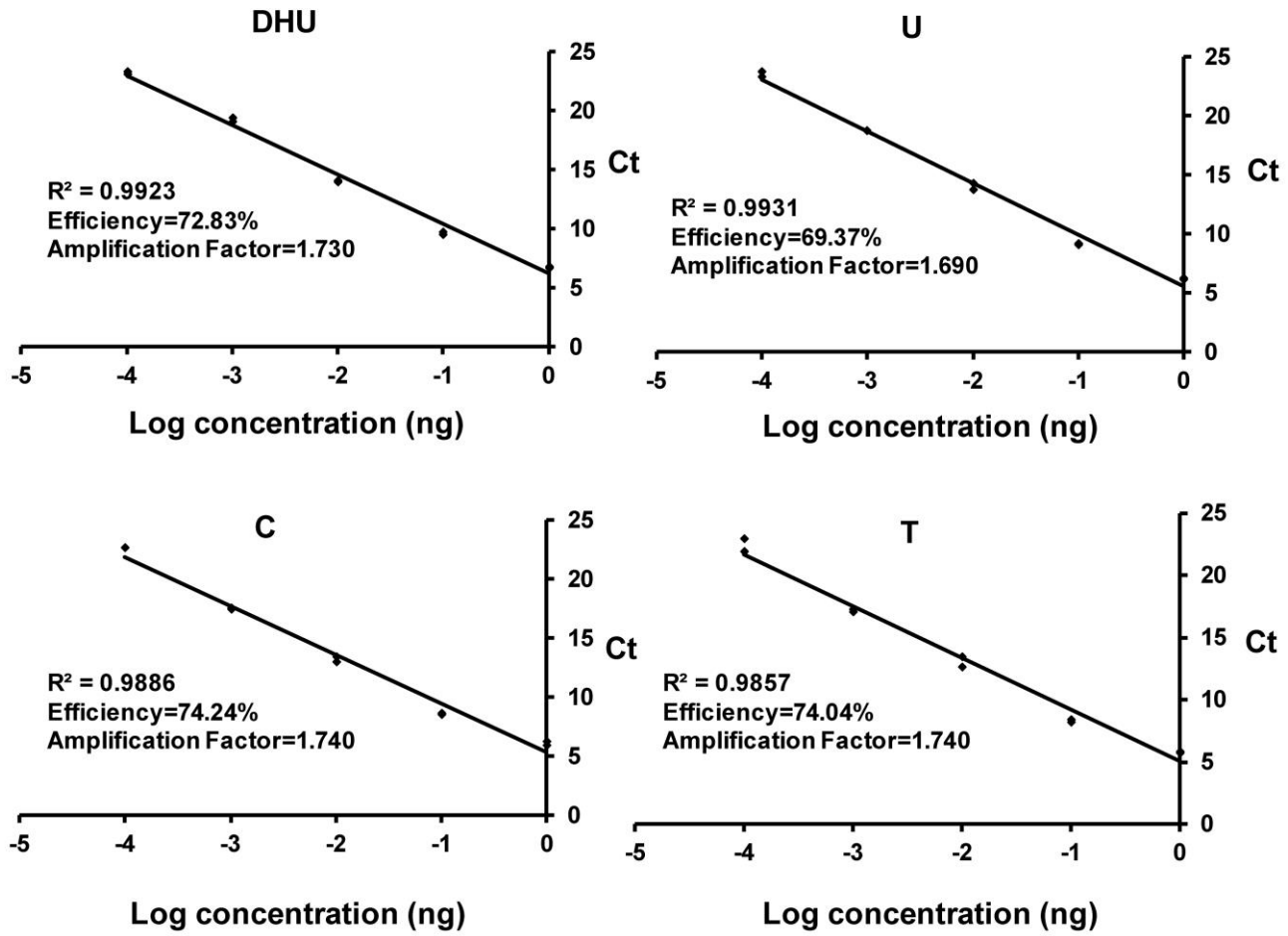
Model DNA containing single methylated and single hydroxymethylated CpG sites was prepared as described in **Supplementary Note 3**. TAPS conversion was done following the **NgTET1 Oxidation** and **Pyridine borane reduction** protocols described in the **Methods**. TAPSβ conversion was done following the **5hmC blocking, NgTET1 Oxidation** and **Pyridine borane reduction** protocols. CAPS conversion was done following the **5hmC oxidation** and **Pyridine borane reduction** protocols. After conversion, 1 ng of converted DNA sample was PCR amplified by Taq DNA Polymerase and processed for Sanger sequencing. TAPS converted both 5mC and 5hmC to T. TAPSβ selectively converted 5mC whereas CAPS selectively converted 5hmC. None of the three methods caused conversion on unmodified cytosine and other bases.



Supplementary Figure 7

TAPS is compatible with various DNA and RNA polymerases and induces complete C-to-T transition shown by Sanger sequencing.

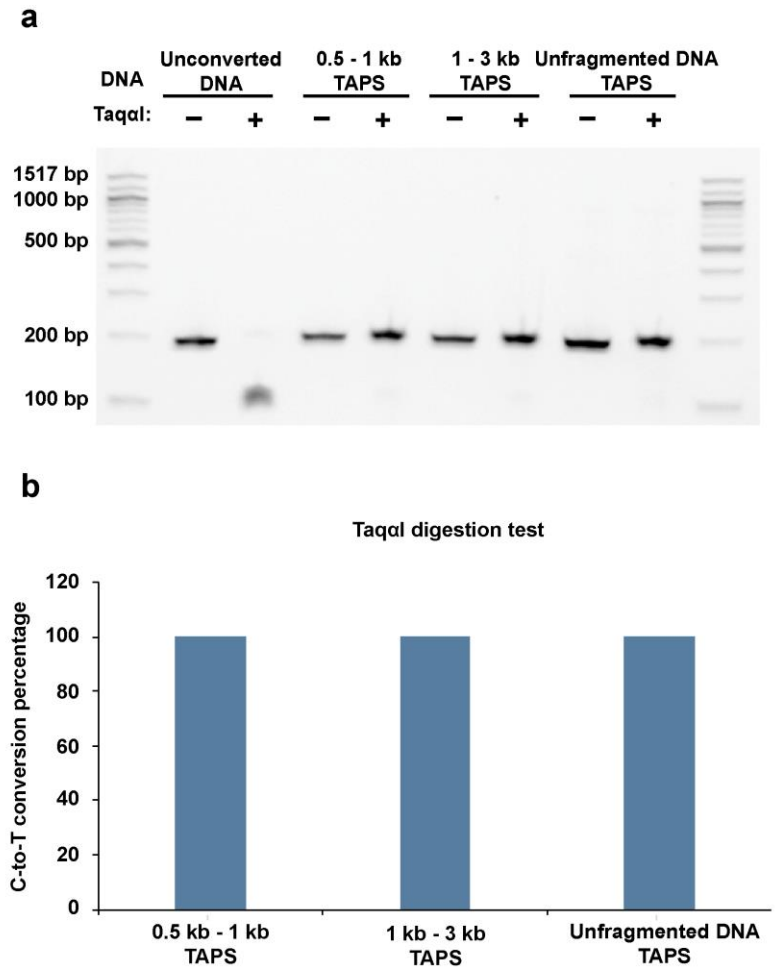
The model DNA containing methylated CpG sites for the polymerase test and primer sequences is described in **Supplementary Note 3**. After TAPS treatment, 5mC was converted to DHU. KAPA HiFi Uracil plus polymerase, Taq polymerase, and Vent exo- polymerase read DHU as T and therefore induce complete C-to-T transition after PCR. Alternatively, primer extension was done with a biotin-labelled primer and isothermal polymerases including Klenow fragment, Bst DNA polymerase, and phi29 DNA polymerase. The newly synthesized DNA strand was separated by Dynabeads® MyOne Streptavidin C1 and then amplified by PCR with Taq polymerase and processed for Sanger sequencing. T7 RNA polymerase could efficiently bypass DHU and insert adenine opposite the DHU site, which is shown by RT-PCR and Sanger sequencing. Other commercial polymerases including KAPA HiFi polymerase, NEB Q5 polymerase, and Phusion polymerase were also tested but failed to amplify DHU containing DNA efficiently.



Supplementary Figure 8

DHU does not show PCR bias compared to T and C.

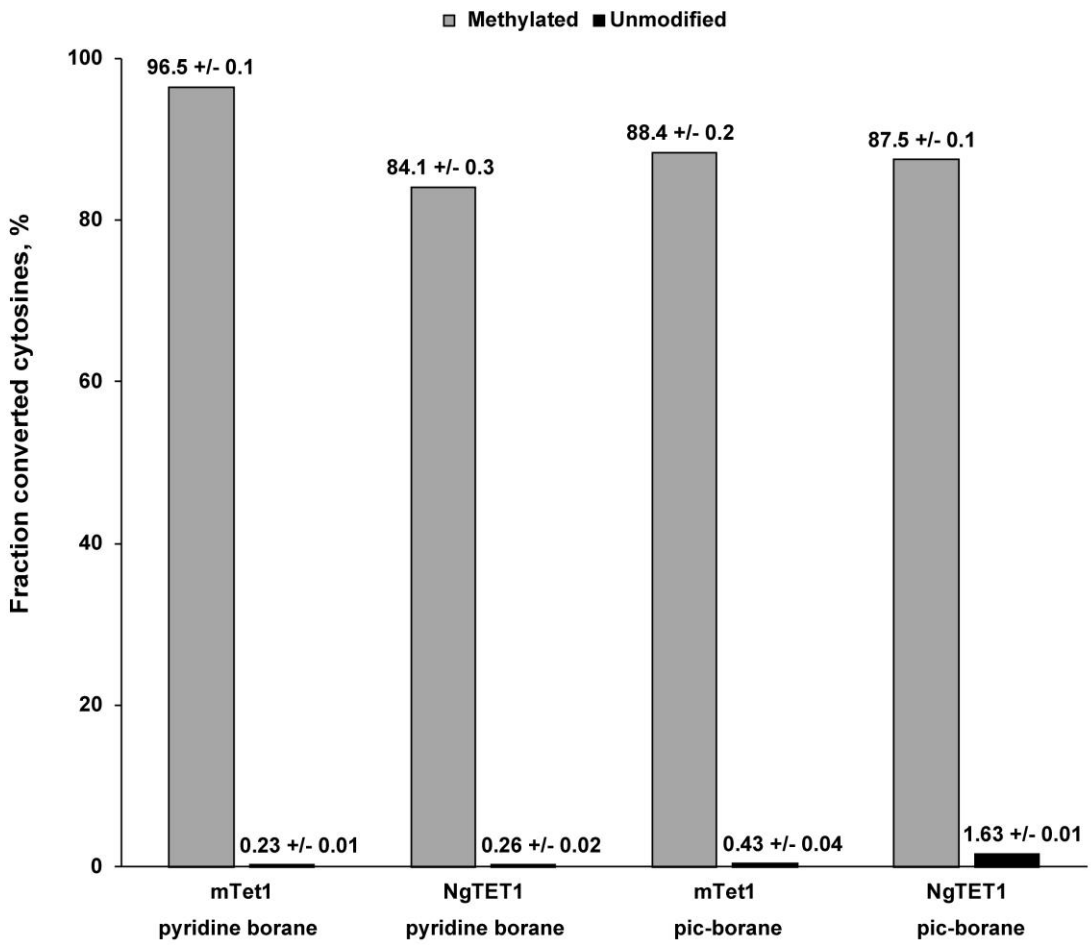
A model DNA containing one DHU/U/T/C modification was synthesized with the corresponding DNA oligos as described in **Supplementary Note 3**. Standard curves for each model DNA with DHU/U/T/C modification were plotted based on qPCR reactions with 1:10 serial dilutions of the model DNA input (from 0.1 pg to 1 ng). Every qPCR experiment was run in triplicates (n=3 technical replicates). The slope of the regression between the log concentration (ng) values and the average Ct values was calculated by SLOPE function in Excel. PCR efficiency was calculated using the following equation:
Efficiency % = (10^(-1/ Slope)-1)*100%
Amplification factor was calculated using the following equation:
Amplification factor=10^(-1/Slope).
The PCR efficiency for the model DNAs with DHU or T or C modification were almost the same, which demonstrated that DHU could be read through as a regular base and would not cause PCR bias.



Supplementary Figure 9

TAPS completely converted 5mC to T regardless of DNA fragment length.

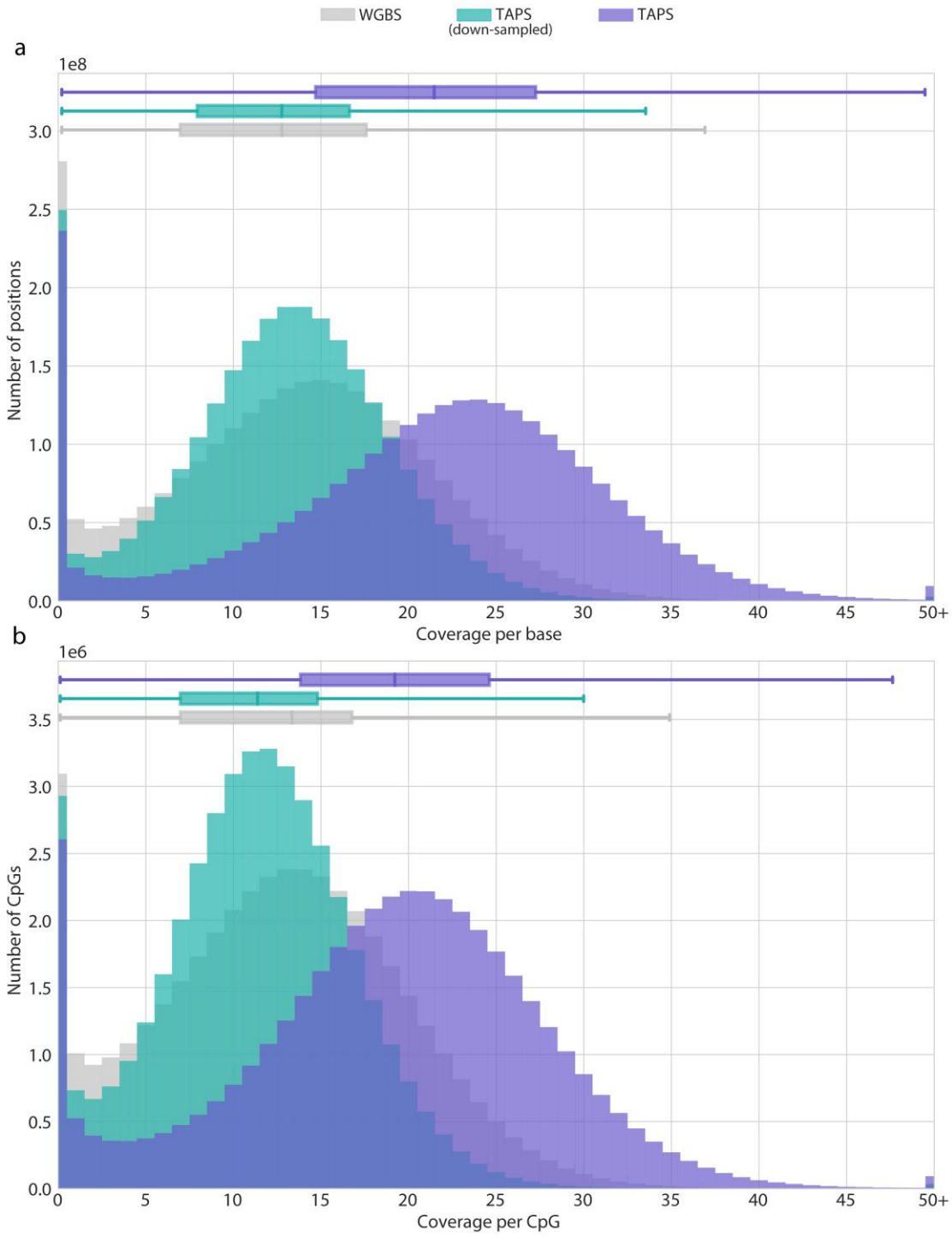
(a) Agarose gel images of the TaqαI-digestion assay confirming complete 5mC to T conversion in all samples regardless of DNA fragment length. 194 bp model sequence from the lambda genome was PCR amplified after TAPS and digested with TaqαI enzyme. The PCR product amplified from unconverted sample could be cleaved, whereas products amplified on TAPS treated samples stayed intact, suggesting loss of restriction site and hence complete C-to-T transition. Experiment was performed once. (b) The C-to-T conversion percentage was estimated by gel band quantification as 100% for all DNA fragment lengths tested.



Supplementary Figure 10

The conversion and false positive rate for different TAPS conditions.

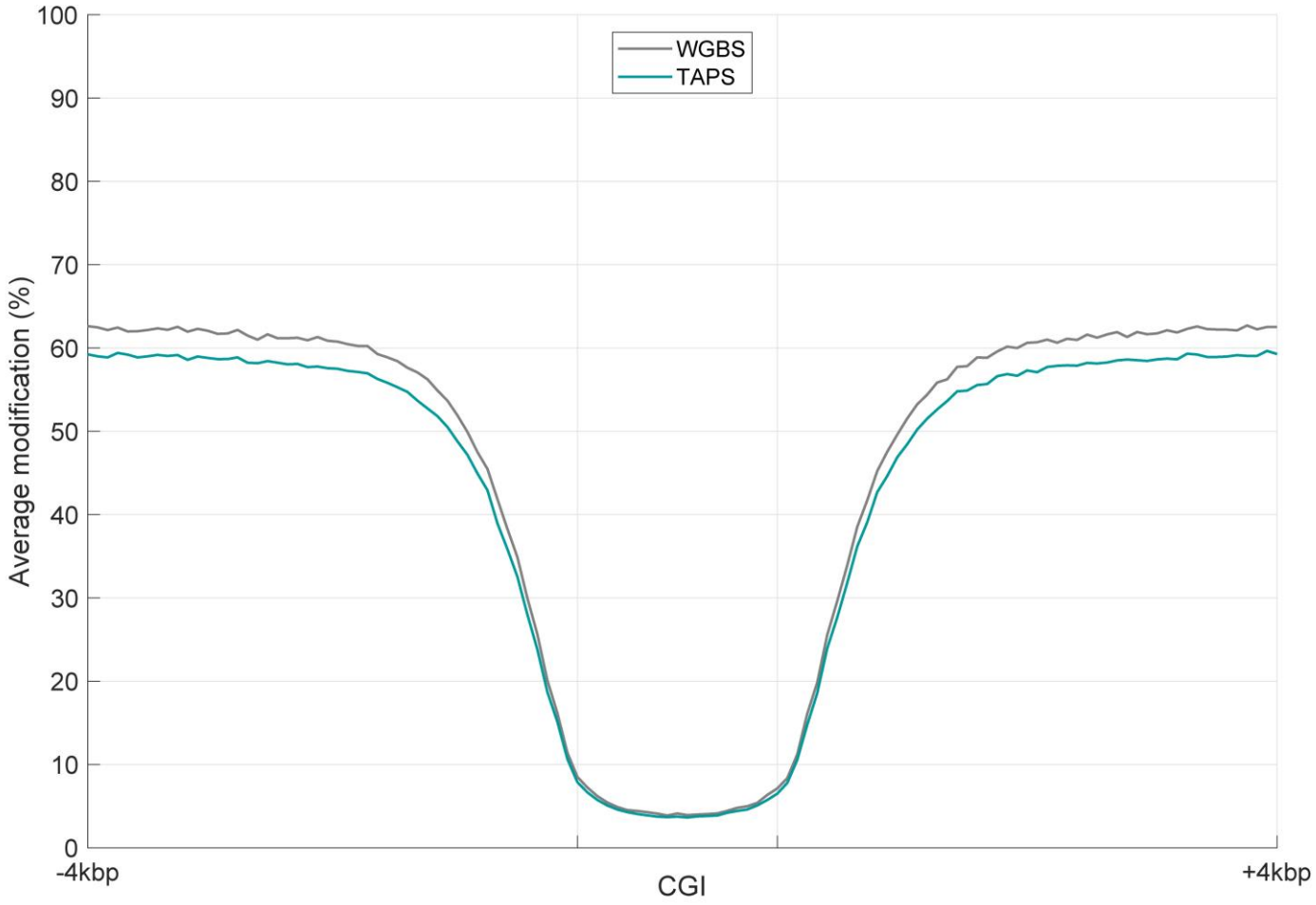
The combination of mTet1 and pyridine borane achieved the highest conversion rate of methylated C (96.5%, calculated with fully CpG methylated Lambda DNA) and the lowest conversion rate of unmodified C (0.23%, calculated with 2kb unmodified spike-in), compared to other conditions with NgTET1 or pic-borane. Shown above are the conversion rates +/- SE of all tested cytosine sites (N of 2kb unmodified positions = 1041, N of covered bacteriophage lambda CpG positions used: mTet1 pyridine borane 6226, mTet1 pic-borane 5871, NgTET1 pyridine borane 5768, NgTET1 pic-borane 6226).



Supplementary Figure 11

TAPS resulted in more even coverage and fewer uncovered positions than WGBS.

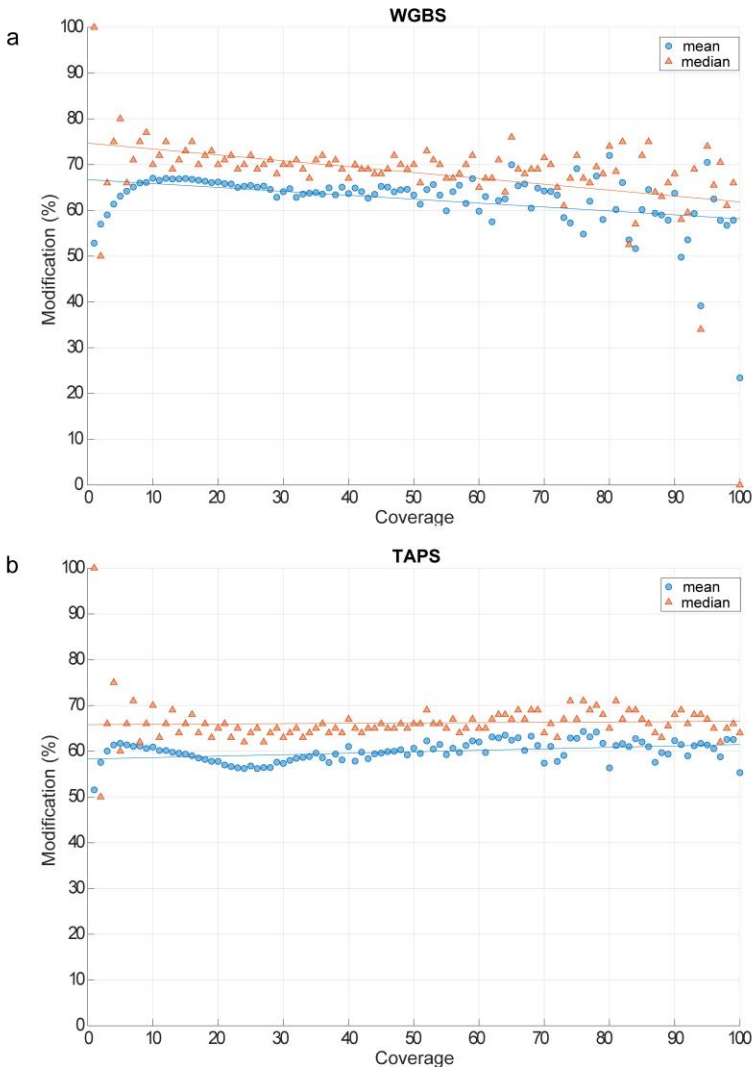
Comparison of coverage depth across (a) all bases (N = 2 725 765 481) and (b) CpG sites (N = 43 445 914, based on mm9 genome, which includes potential genetic variants in E14 genome) between WGBS and TAPS, computed on both strands. For “TAPS (down-sampled)”, random reads out of all mapped TAPS reads were selected so that the median coverage matched the median coverage of WGBS. Positions with coverage above 50× are shown in the last bin.



Supplementary Figure 12

Modification levels around CpG Islands.

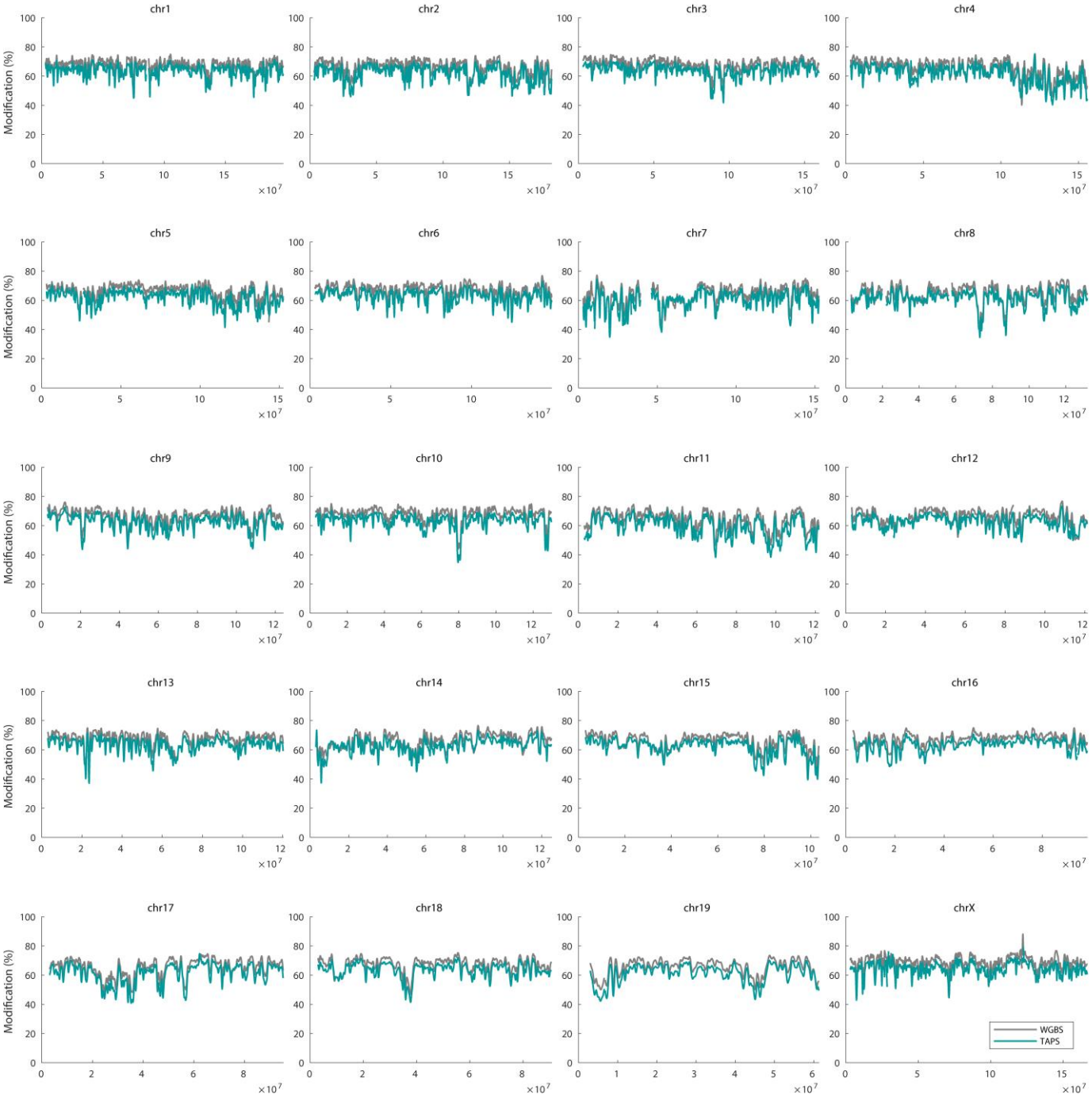
Average modification levels in CpG islands (binned into 20 windows) and 4 kb flanking regions (binned into 50 equally sized windows). Bins with coverage below 3 reads were ignored.



Supplementary Figure 13

TAPS exhibits smaller coverage-modification bias than WGBS.

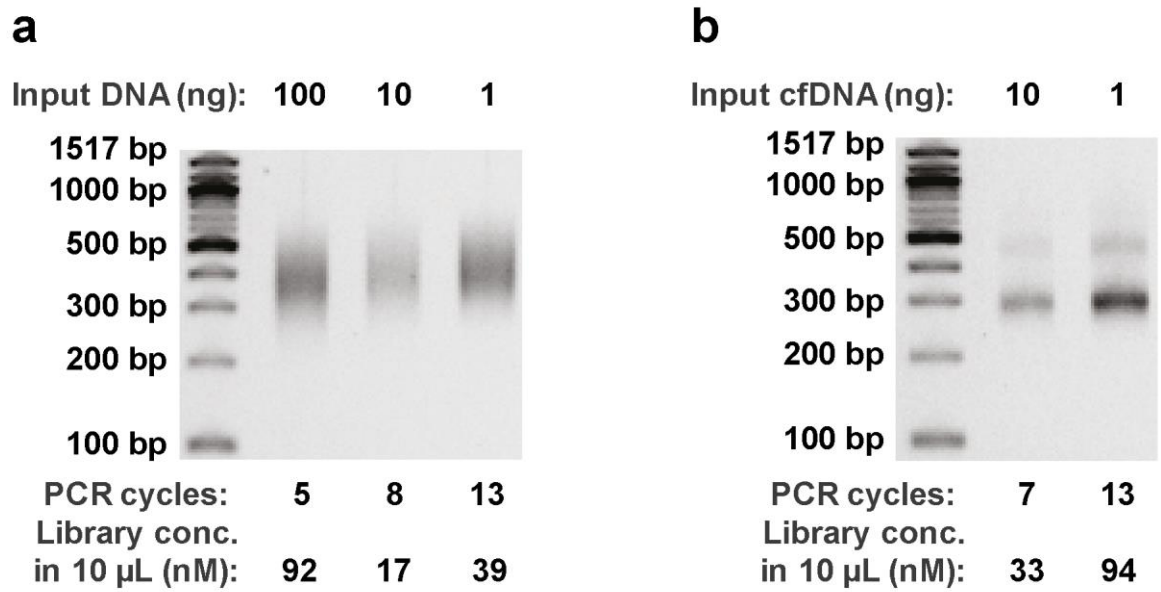
All CpG sites were binned according to their coverage, and the mean (blue) and the median (orange) modification values are shown in each bin for WGBS (**a**) and TAPS (**b**). The CpG sites covered by more than 100 reads are shown in the last bin. The lines represent a linear fit through the data points.



Supplementary Figure 14

Distribution of modification levels across all chromosomes.

Average modification levels in 100 kb windows along mouse chromosomes, weighted by the coverage of CpG, and smoothed using a Gaussian weighted moving average filter with window size 10.



Supplementary Figure 15

Low-input gDNA and cell-free DNA TAPS libraries prepared with dsDNA KAPA HyperPrep library preparation kit.

Sequencing libraries were successfully constructed with as little as 1 ng of (**a**) mESC gDNA and (**b**) cell-free DNA with KAPA HyperPrep kit. Experiment was performed once. Note that cell-free DNA has a sharp length distribution around 160 bp (nucleosome size) due to plasma nuclease digestion. After library construction, it becomes ~300bp, which is the sharp band in (**b**).

Supplementary Table 1.

Comparison of bisulfite sequencing (BS) and related methods versus Pyridine borane sequencing (PS) for 5fC and 5caC sequencing.

Base	BS	fCAB-Seq /redBS-Seq	caCAB-Seq	fC-CET	**PS**	**PS with 5fC blocking**	**PS with 5caC blocking**
C	T	T	T	C	**C**	**C**	**C**
5mC	C	C	C	C	**C**	**C**	**C**
5hmC	C	C	C	C	**C**	**C**	**C**
5fC	T	C	T	T	**T**	**C**	**T**
5caC	T	T	C	C	**T**	**T**	**C**

Supplementary Table 2.

Mapping and sequencing quality statistics for WGBS and TAPS.

Measure	WGBS	TAPS
Total raw reads	376,062,375	455,548,210
Trimmed reads	367,860,813	453,028,186
Mapped reads (mm9+spike-ins+PhiX)	251,940,139	451,077,132
PCR deduplicated reads	232,303,596	398,127,851
Mapping rate (mapped reads/trimmed reads)	68.49%	99.57%
Unique mapping rate (unique reads [MAPQ>0 for TAPS]/trimmed reads)	68.49%	88.08%
Unique PCR deduplicated mapping rate (unique PCR deduplicated reads [MAPQ>0 for TAPS] /trimmed reads)	63.15%	81.31%

Supplementary Table 3.

Coverage statistics for TAPS, WGBS and TAPS down-sampled to have approximately the same mean coverage as WGBS. Here, coverage was computed for both strands at all positions in the genome.

Measure	WGBS	TAPS with down-sampling	TAPS without down-sampling
Mean	13.078	12.411	21.001
Variance	1988.242	482.242	1371.912
median	13	13	22
qtl25	7	8	15
qtl75	18	17	28
iqr	11	9	13
maximum	116084	37329	63526

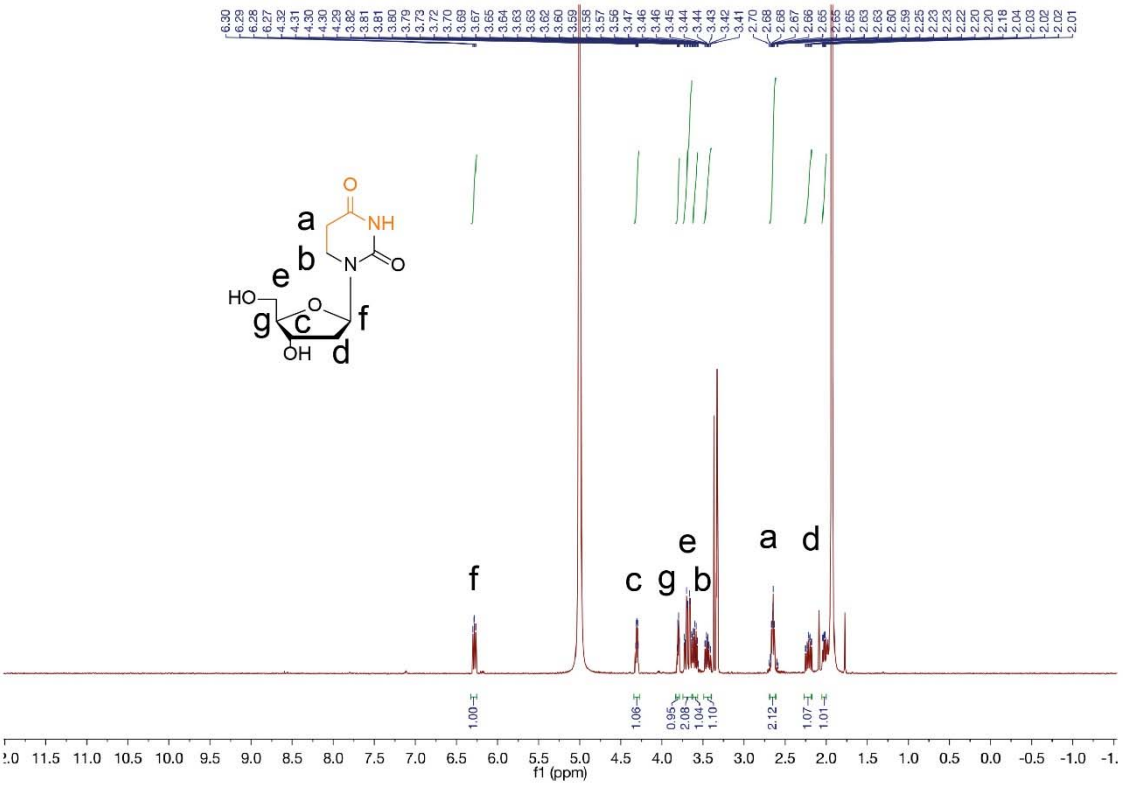
Supplementary Table 4.

Compound-dependent HPLC-MS/MS parameters used for nucleosides quantification. RT: retention time, CE: collision energy; CAE: cell accelerator voltage. All the nucleosides were analyzed in the positive mode.

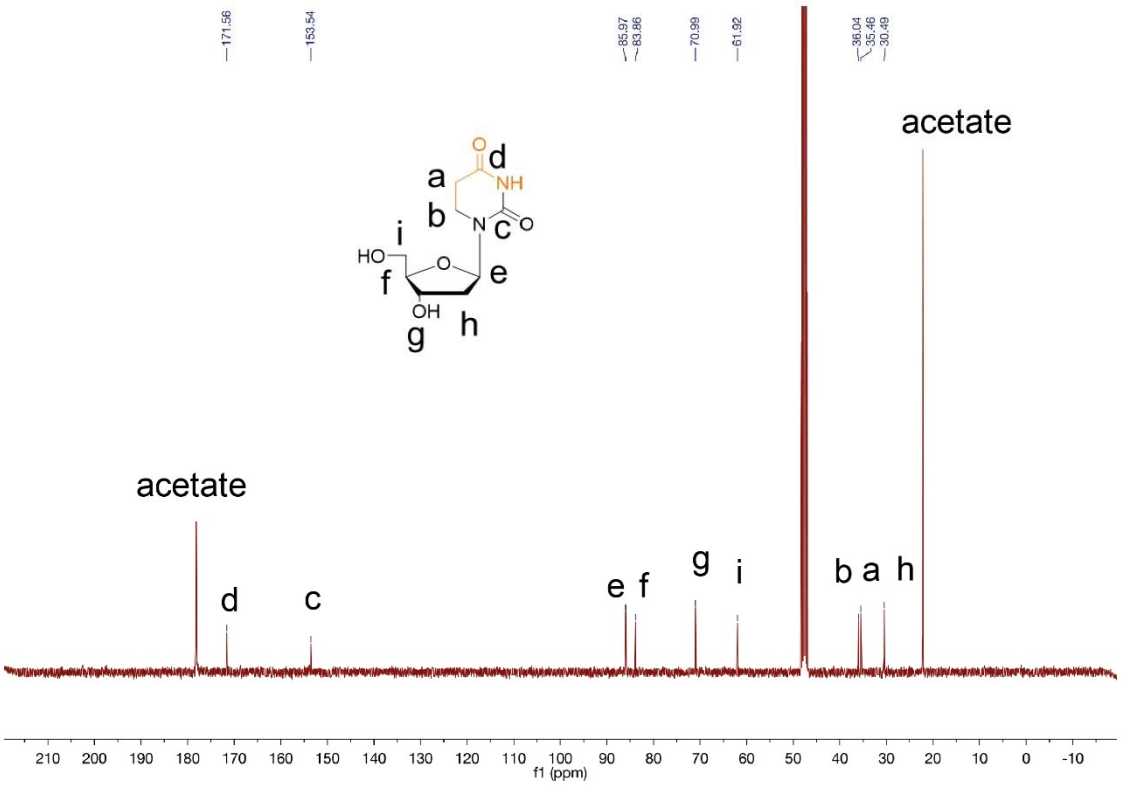
Compound	Precursor Ion (m/z)	Product Ion (m/z)	RT (min)	Delta RT(min)	CE (V)	CAE (V)
dA+H	252	136	13.78	2	10	4
dT+H	243	127	11.07	2	10	4
dT+Na	265	149	11.07	2	10	4
dG+H	268	152	9.64	2	10	4
dC+H	228	112	3.71	1.5	10	4
dC+Na	250	134	3.71	1.5	10	4
mdC+H	242	126	9.05	1.5	10	4
mdC+Na	264	148	9.05	1.5	10	4
hmdC+H	258	142	4.34	2	12	4
hmdC+Na	280	164	4.34	2	12	4
fdC+H	256	140	10.69	2	8	4
fdC+Na	278	162	10.69	2	8	4
cadC+H	272	156	1.75	3	12	4
cadC+Na	294	178	1.75	3	12	4
DHU+H	231	115	3.45	3	10	4
DHU+Na	253	137	3.45	3	10	4

Supplementary Note 1: Single nucleoside pic-borane reaction

500 µL of 3.26 M 2-picoline-borane (pic-borane, Sigma-Aldrich) in MeOH and 500 µL of 3 M sodium acetate solution (pH 5.2, Thermo Fisher) were added into 10 mg of 2'-deoxycytidine-5-carboxyllic acid sodium salt (Berry&Associates). The mixture was stirred for 1 hour at 60°C. The product was purified by HPLC to give a pure compound as white foam. High resolution MS (Q-TOF) *m/z* $[M + Na]^+$ calculated for $C_9H_{14}N_2O_5Na$: 253.0800; Found: 253.0789. All experiments were performed once.



^{1}H NMR (MeOH-d_4, 400 MHz) chart of the single nucleoside pic-borane reaction product. δ ppm: 6.28 (t, 1H, J = 7 Hz), 4.30 (m, 1H), 3.81 (m, 1H), 3.63 (m, 2H), 3.46 (m, 2H), 2.65 (t, 2H, J = 6 Hz), 2.20 (m, 1H), 2.03 (m, 1H).



^{13}C NMR (MeOH-d_4, 400 MHz) chart of the single nucleoside pic-borane reaction product. δ ppm: 171.56 (CO), 153.54 (CO), 85.97 (CH), 83.86 (CH), 70.99 (CH), 61.92 (CH_2), 36.04 (CH_2), 35.46 (CH_2), 30.49 (CH_2).

Supplementary Note 2: Preparation of model DNA

1. DNA oligos for MALDI and HPLC-MS/MS test

Regular DNA oligos and modified DNA oligos with 5mC or 5hmC were purchased from IDT. All the sequences and modifications can be found in **Supplementary Fig. 3** and **4**.

The DNA oligo with 5fC modification was synthesized by the C-tailing method. DNA oligos 5'-GTCGACCGGATC-3' and 5'-TTGGATCCGGTCGACTT-3' were annealed and then incubated with 5-formyl-2'-dCTP (Trilink Biotech) and Klenow Fragment 3'→5' exo- (NEB) in NEBuffer 2 for 2 hours at 37°C. The product was purified with Bio-Spin® P-6 Gel Columns (Bio-rad).

The DNA oligo with 5caC modification was synthesized using Expedite 8900 Nucleic Acid Synthesis System (Thermo Fisher) with standard phosphoramidites (Sigma) and 5-Carboxy-dC-CE Phosphoramidite (Glen Research). Subsequent deprotection and purification were carried out with Glen-Pak Cartridges (Glen Research) according to the manufacturer's instructions. Purified oligonucleotides were characterized by Voyager-DE MALDI-TOF Biospectrometry Workstation.

2. 222 bp Model DNA for conversion test

The 222 bp model DNA containing five CpG sites was generated from lambda DNA (Thermo Fisher) by PCR amplification using Taq DNA Polymerase (NEB) and purified by AMPure XP beads (Beckman Coulter). Primers sequences are as follows: FW-5'-CCTGATGAAACAAGCATGTC-3', RV-5'-CAUTACTCACUTCCCCACUT-3'. The uracil base in the reverse strand of PCR product was removed by USER enzyme (NEB). 100 ng of purified PCR product was then methylated in 20 µl solution containing 1X NEBuffer 2, 0.64 mM S-adenosylmethionine and 20 U M.SssI CpG Methyltransferase (NEB) for 2 h at 37°C, followed by 20 min heat inactivation at 65°C. The methylated 222 bp model DNA was purified by AMPure XP beads.

3. Model DNA for TAPS, TAPSβ and CAPS validation with Sanger sequencing

The 34 bp DNA oligo containing single 5mC and single 5hmC sites was annealed with other DNA oligos in the annealing buffer containing 5 mM Tris-Cl (pH 7.5), 5 mM $MgCl_2$, and 50 mM NaCl, and then ligated in a reaction containing 400 U T4 ligase (NEB) at 25°C for 1 h and purified by 1.8X AMPure XP beads.

DNA	Sequence (5' to 3')
34 bp mC and hmC	CCCGA^{m}**C**GCATGATCTGTACTTGATCGAC^{hm}**C**GTGCAAC
TruSeq Universal Adapter	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
TruSeq Adapter (Index 6)	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCACGCCAATATCTCGTATGCCGTCTTCTGCTTG
Uracil linker	TCTTCCGAUCGTTGCACGGUCGATCAAGUACAGATCATGCGUCGGGAGAUCGGAAG

The uracil linker was removed by USER enzyme after ligation reaction.

The final product sequence (5' to 3'):

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCCGA^{m}**C**GCATGATCTGTACTTGATCGAC^{hm}**C**GTGCAACGATCGGAAGAGCACACGTCTGAACTCCAGTCACGCCAATATCTCGTATGCCGTCTTCTGCTTG

The PCR primer to amplify the model DNA:

P5: 5'-AATGATACGGCGACCACCGAG-3'

P7: 5'-CAAGCAGAAGACGGCATACGAG-3'

4. Model DNA for polymerase test and Sanger sequencing

The model DNA for polymerase test and Sanger sequencing was prepared with the same ligation method above except different DNA oligos were used:

DNA	Sequence (5' to 3')
34 bp mC	AGCAGTCT^{m}**C**GATCAGCTG^{m}**C**TACTGTA^{m}**C**GTAGCAT
TruSeq Universal Adapter	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
TruSeq Adapter (Index 6)	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCACGCCAATATCTCGTATGCCGTCTTCTGCTTG
Insert_1_40_bp	/5Phos/AGGTGCGCTAAGTTCTAGATCGCCAACTGGTTGTGGCCTT
Insert_2_60_bp	/5Phos/CTATAGCCGGCTTGCTCTCTCTGCCTCTAGCAGCTGCTCCCTATAGTGAGTCGTATTAAC
40_bp-Linker-1	ATCTAGAACTTAGCGCACCTAGATCGGAAGAGCGTCGTGT
80_bp-Linker:	AGAGAGCAAGCCGGCTATAGATGCTACGTACAGTAGCAGCTGATCAAGACTGCTAAGGCCACAACCAGTTGGCG
42_bp-Linker-2:	AGACGTGTGCTCTTCCGATCGTTAATACGACTCACTATAGGG

Final product sequence (5' to 3'):

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGGTGCGCTAAGTTCTAGATCGCCAACTGGTTGTGGCCTTAGCAGTCT^{m}**C**GATCAGCTG^{m}**C**TACTGTA^{m}**C**GTAGCATCTATAGCCG

GCTTGCTCTCTCTGCCTCTAGCAGCTGCTCCCTATAGTGAGTCGTATTAACGATCGGAAGAGCACACGTCTG
AACTCCAGTCACGCCAATATCTCGTATGCCGTCTTCTGCTTG
PCR primer to amplify the model DNA:
P5: 5'-AATGATACGGCGACCACCGAG-3'
P7: 5'-CAAGCAGAAGACGGCATACGAG-3'
Biotin-labelled primer sequence for primer extension:
Biotin-P7: 5'-Bioin-CAAGCAGAAGACGGCATACGAG-3'
PCR primer for RT-PCR after T7 RNA polymerase transcription:
P5: 5'-AATGATACGGCGACCACCGAG-3'
RT: 5'-TGCTAGAGGCAGAGAGAGCAAG-3'

5. Model DNA for PCR bias test

The model DNA for PCR bias test was prepared with the same ligation method above except different DNA oligos were used:

DNA	Sequence (5' to 3')
17 bp X	AGCAGTCT**X**GATCAGCT
17 bp No Modification	GCTACTGTACGTAGCAT
TruSeq Universal Adapter	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTC CGATCT
TruSeq Adapter (Index 6)	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCACGCCAATATCTCGTA TGCCGTCTTCTGCTTG
Insert_1_40_bp	/5Phos/AGGTGCGCTAAGTTCTAGATCGCCAACTGGTTGTGGCCTT
Insert_2_60_bp	/5Phos/CTATAGCCGGCTTGCTCTCTCTGCCTCTAGCAGCTGCTCCCTATAGT GAGTCGTATTAAC
40_bp-Linker-1	ATCTAGAACTTAGCGCACCTAGATCGGAAGAGCGTCGTGT
80_bp-Linker:	AGAGAGCAAGCCGGCTATAGATGCTACGTACAGTAGCAGCTGATCAAGACT GCTAAGGCCACAACCAGTTGGCG
42_bp-Linker-2:	AGACGTGTGCTCTTCCGATCGTTAATACGACTCACTATAGGG

Final product sequence (5' to 3'):
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGGTGCGCTAAGTT
CTAGATCGCCAACTGGTTGTGGCCTTAGCAGTCT**X**GATCAGCTGCTACTGTACGTAGCATCTATAGCCGGCT
TGCTCTCTCTGCCTCTAGCAGCTGCTCCCTATAGTGAGTCGTATTAACGATCGGAAGAGCACACGTCTGAAC
TCCAGTCACGCCAATATCTCGTATGCCGTCTTCTGCTTG
X= DHU or U or T or C.
PCR primer to amplify the model DNA:
P5: 5'-AATGATACGGCGACCACCGAG-3'
P7: 5'-CAAGCAGAAGACGGCATACGAG-3'

Supplementary Note 3: Preparation and sequences of spike-in DNAs used in sequencing

1. CpG methylated lambda DNA

1 µg of unmethylated lambda DNA (Promega) was methylated in the presence of 0.64 mM SAM and 0.8 U/µl M.SssI enzyme (NEB) in Mg^{2+}-free buffer (10 mM Tris-Cl pH 8.0, 50 mM NaCl, and 10 mM EDTA) for 2 h at 37°C in a 50 µL reaction. Then, 0.5 µL of M.SssI enzyme and 1 µL of SAM were added and the reaction was incubated for an additional 2 h at 37°C. Methylated DNA was subsequently purified on 1X Ampure XP beads. To assure complete methylation, whole procedure was repeated a second time in NEBuffer 2. DNA methylation was then validated with HpaII digestion assay. 50 ng of methylated and unmethylated DNA were digested in a 10 µL reaction with 2 U of HpaII enzyme (NEB) in CutSmart buffer (NEB) for 1 h at 37°C. Digestion products were run on a 1% agarose gel together with undigested lambda DNA control. Unmethylated lambda DNA was digested after the assay whereas methylated lambda DNA remained intact confirming complete and successful CpG methylation. The sequence of lambda DNA can be found in GenBank - EMBL Accession Number: J02459.

2. 2 kb unmodified spike-in control

The 2 kb unmodified spike-in control was PCR amplified from the pNIC28-Bsa4 plasmid (Addgene, cat. no. 26103) in the reaction containing 1 ng DNA template, 0.5 µM primers, 1 U Phusion High-Fidelity DNA Polymerase (Thermo Fisher). Primers sequences are as follows: FW-5'- CACAGATGTCTGCCTGTTCA -3', RV-5'- AGGGTGGTGAATGTGAAACC -3'. PCR product was purified on Zymo-Spin column. 2 kb unmodified control sequence (5' to 3'):

CACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGA
TAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGT
TCATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCC
CGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCA
GGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGC
AGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAACCGAAG
ACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGT
GATTCATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCAT
GCGCACCCGTGGGGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCA
GTGACGAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGC
TCCAGCGAAAGCGGTCCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATGATA
AAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCGCCCACCGGAAGGAGCTGACTGGGTTGA
AGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCT
CACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAG
AGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCC
TTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCT
GTTTGATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATAT
CCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAAC
CAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCA
GTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCA
GACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATG
CTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATC
AAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGT
TAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTT
CGTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTG
CGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGT
TGTGCCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAA
ACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTA
TAACGTTACTGGTTTCACATTCACCACCCT

3. Synthetic spike-in with N5mCNN and N5hmCNN

A synthetic oligo with N5mCNN and N5hmCNN sequences was produced by the annealing and extension method. Oligo sequences are listed below.

Oligo	Template sequence (5' to 3')
N5mCNN	GAAGATGCAGAAGACAGGAAGGATGAAACACTCAGGCGCACGCTGGCATN^{m}**C**NNGACAAACCACAAGAACAGGCTAGTGAGAATGAAGGGA
N5hmCNN	CCAACTCTGAAACCCACCAACGCCAACATCCACCACACAACCCAAGATN^{hm}**C**NNGACCATCTTACAAACATATCCCTTCATTCTCACTAGCC

Briefly, 10 µM N5mCNN and N5hmCNN oligos (IDT) were annealed in the annealing buffer containing 5 mM Tris-Cl (pH 7.5), 5 mM $MgCl_2$, and 50 mM NaCl. Extension was performed in the NEB buffer 2 with 0.4 mM dNTPs (dATP/dGTP/dTTP/dCTP) and 5 U of Klenow Polymerase (NEB) for 1 hour at 37°C. After the reaction, the spike-in control was purified on Zymo-Spin column (Zymo Research).

Synthetic spike-in with N5mCNN and N5hmCNN (5' to 3'):

GAAGATGCAGAAGACAGGAAGGATGAAACACTCAGGCGCACGCTGGCATN^{m}**C**NNGACAAACCACAAGAACA
GGCTAGTGAGAATGAAGGGATATGTTTGTAAGATGGTCNNGNATCTTGGGTTGTGTGGTGGATGTTGGCGT
TGGTGGGTTTCAGAGTTGG

Complementary strand (5' to 3'):

CCAACTCTGAAACCCACCAACGCCAACATCCACCACACAACCCAAGATN^{hm}**C**NN
GACCATCTTACAAACATATCCCTTCATTCTCACTAGCCTGTTCTTGTGGTTTGTCNNGNATGCCAGCGTGCG
CCTGAGTGTTTCATCCTTCCTGTCTTCTGCATCTTC

Supplementary Note 4

Sequence of mTet1CD insert (NM_001253857.2, 4371-6392) with N-terminal Flag-tag is cloned into pcDNA3-Flag between KpnI and BamH1 restriction sites.

ATGGACTACAAAGACGATGACGACAAGGAAGCTGCACCCTGTGACTGTGATGGAGGTACACAAAAAGAAAA
AGGCCCATATTATACACACCTTGGGGCAGGACCAAGTGTGGCTGCTGTCAGGGAGCTCATGGAGACTAGGT
TTGGCCAGAAGGGGAAGGCAATCCGGATTGAGAAGATAGTGTTCACGGGGAAGGAAGGGAAGAGCTCTCA
GGGCTGCCCGGTCGCCAAGTGGGTGATCAGAAGAAGTGGTCCTGAAGAGAAGCTTATTTGTTTGGTTCGTG
AGCGTGTAGACCATCACTGTTCGACGGCTGTGATAGTTGTCCTTATCCTGCTGTGGGAAGGTATCCCTCGCC
TGATGGCTGACCGCCTGTACAAAGAGCTCACTGAGAACTTGAGGTCCTACAGCGGACATCCCACAGACCGA
AGATGTACCCTCAACAAAAAGCGTACCTGCACCTGTCAAGGCATCGACCCAAAAACCTGCGGAGCGTCCTT
CTCCTTTGGCTGTTCGTGGAGCATGTATTTCAACGGCTGTAAGTTTGGGAGGAGTGAAAACCCCAGAAAATT
CAGACTTGCTCCAAACTACCCCTTACATAACTACTATAAGAGAATTACTGGAATGAGTTCTGAAGGAAGTGAC
GTGAAAACCGGGTGGATCATTCCAGACCGCAAGACCCTCATAAGCAGAGAGGAAAAACAGCTTGAAAAGAA
TTTACAAGAATTGGCTACAGTATTAGCTCCACTTTACAAGCAGATGGCTCCAGTTGCTTATCAAAATCAGGTG
GAATATGAAGAAGTTGCTGGAGACTGTCGACTTGGAAATGAAGAGGGGCGTCCTTTCTCTGGTGTCACCTGT
TGCATGGATTTTTGTGCCCATTCTCACAAGGACATTCACAACATGCACAACGGAAGCACCGTGGTGTGTACG
TTGATTCGAGCAGATGGCCGTGACACAAATTGTCCCGAGGATGAACAACTCCACGTCCTGCCACTATACCG
GCTTGCAGACACTGATGAATTTGGCTCCGTGGAAGGGATGAAGGCCAAAATCAAATCTGGGGCCATCCAAG
TCAATGGGCCAACCAGGAAGAGGCGACTACGTTTTACTGAGCCTGTTCCTCGATGTGGGAAGAGGGCCAAA
ATGAAGCAGAACCACAATAAATCAGGTTCACACAACACTAAGAGCTTTTCATCAGCCTCATCTACTTCTCACC
TAGTGAAAGACGAATCTACAGACTTCTGTCCCCTGCAGGCTTCCTCCGCAGAAACATCTACCTGTACGTACA
GTAAAACAGCCTCAGGTGGGTTTGCAGAAACAAGTAGTATTCTCCACTGCACAATGCCTTCTGGAGCACACA
GTGGTGCTAATGCAGCTGCTGGGGAATGTACTGGAACGGTGCAGCCTGCCGAGGTGGCTGCTCATCCTCA
CCAGTCTCTTCCCACAGCCGATTCTCCCGTTCATGCTGAGCCTCTCACTAGTCCATCTGAGCAGCTAACTTC
TAACCAGTCAAACCAGCAGCTCCCTCTCCTCAGCAATTCTCAGAAACTGGCTTCCTGTCAGGTGGAAGATGA
GCGGCACCCTGAAGCGGATGAGCCTCAGCACCCCGAGGACGATAACTTGCCTCAACTTGATGAATTCTGGT
CAGACAGTGAGGAGATCTACGCCGATCCTTCCTTTGGTGGCGTGGCGATAGCACCCATTCACGGCTCGGTG
CTCATTGAGTGCGCTCGGAAGGAGCTTCATGCTACCACCTCTTTGCGCTCCCCCAAACGAGGGGTCCCTTT
TCGTGTGTCCCTTGTATTCTACCAGCACAAAAGCCTAAACAAGCCTAATCATGGTTTTGATATCAACAAAATT
AAGTGTAAATGCAAAAAAGTAACGAAAAAAAAGCCCGCAGACCGGGAGTGTCCTGATGTATCCCCCGAAGC
CAATTTATCACACCAAATTCCTTCTCGAGTTGCATCAACCTTAACCCGAGACAATGTTGTTACCGTGTCCCCA
TACTCTCTCACTCATGTTGCGGGACCCTACAATCGTTGGGTCTAA

Supplementary Note 5: Production of recombinant mTet1CD

Introduction

mTet1CD protein is expressed as an N-terminal Flag-tagged protein in Expi293F™ (Gibco) cells. The protocol is optimised for suspension cells and produces up to 0.5 mg of protein from 1 L of culture. Application of mammalian expression system improves protein stability and activity while Flag-tag purification enables isolation of homogenous protein after a single purification step. The activity and quality of the recombinant mTet1CD is checked by MALDI Mass Spectrometry analysis. Based on this assay, recombinant mTet1CD is fully active and able to catalyse oxidation of 5mC to 5caC. Any significant digestion of the tested model oligo was detected by MALDI, confirming that protein is free from nucleases. The detailed protocol for the expression and purification of mTet1CD enzyme is described below.

Reagents:

- Expi293F™ Cells (Gibco, A14527)
- FreeStyle™ 293 Expression Medium (Gibco, 12338018)
- 2 l cell culture roller bottles (e.g. CELLMASTER™, 680068)
- Polyethylenimine (Polysciences, 23966-1)
- DPBS, no calcium, no magnesium (Gibco, 14190-094)
- UltraPure™ 1 M Tris-HCl Buffer, pH 7.5 (Invitrogen, 15567027)
- 5M NaCl (Invitrogen, AM9759)
- Triton X-100 (Sigma, T8787)
- cOmplete™ Protease Inhibitor Cocktail EDTA-free (Sigma, 11873580001)
- PMSF (ThermoScientific, 36978)
- 3X Flag Peptide (Sigma, F4799)
- HEPES (Sigma, H3375)
- DTT (Fluorochem, M02712)
- ANTI-FLAG® M2 Affinity Gel (Sigma, A2220)
- Empty gravity chromatography column (e.g. Econo-Pac® Chromatography Column, Bio-rad, 7321010)
- 3X Flag Peptide (Sigma, F4799)
- Glycine (Sigma, G7126)
- Amicon® Ultra-4 Centrifugal Filter Unit MWCO = 30 kD (Merck, UFC803024)
- Bio-Spin® P-30 Gel Columns, Tris Buffer (Bio-rad, 7326231)

mTet1CD expression

1. Seed Expi293F™ Cells at 0.5 x 10^6 cells/ml density into a final volume of 1000 ml FreeStyle™ 293 Expression Medium in a 2 L roller bottle and grow at 37°C, 170 rpm, and 5% CO_2 until cells reach a density of 1.0 x 10^6 cells/mL. It takes approximately 24 hours.
2. Transfect cells with the pcDNA3-Flag-mTet1CD plasmid:
   a) Use 1 µg of pCDNA3-Flag-mTet1CD plasmid to transfect 1 million cells. For 1 L culture at density 1.0 x 10^6 cells/ml, dilute 1 mg of plasmid in 100 mL PBS and add 3 mL Polyethylenimine (PEI, stock concentration 1 mg/mL). Final DNA to PEI ratio is 1:3. Vortex DNA:PEI solution for at least 10 s and incubate at room temperature for 20 min.
   b) Slowly add DNA:PEI solution to 1 L cell culture and mix well. Maintain cells at 37°C, 170 rpm, and 5% CO_2 for 48 h.
3. After 48 h, collect cells by centrifugation at 3000 x g and 4˚C for 20 min. Remove supernatant and proceed with purification protocol. Alternatively, the pellet can be stored at -80°C for few days without affecting protein activity.

mTet1CD purification

Buffers

Lysis Buffer: 50 mM Tris-Cl pH = 7.5, 500 mM NaCl, 1X cOmplete™ Protease Inhibitor Cocktail, 1 mM PMSF, 1% Triton X-100.

TBS: 50 mM Tris pH = 7.5, 150 mM NaCl.

Elution buffer: 20 mM HEPES pH = 8.0, 150 mM NaCl, 0.1 mg/mL 3X Flag, 1X cOmplete™ Protease Inhibitor Cocktail, 1 mM PMSF.

Storage buffer: 20 mM HEPES pH = 8.0, 150 mM NaCl, 1 mM DTT.

1. Resuspend 1 L cell pellet in 100 mL lysis buffer. Incubate on ice for 20 min.
   Note: Volume of lysis buffer can be scaled-up if cell pellets from more than 1 L culture are purified.
2. Clarify the lysate by centrifuging at 30000 x g and 4°C for 30 min. Collect clear supernatant to clean glass bottle. Save 5 µL for SDS-PAGE gel.
   Note: If larger lysate volumes are used, centrifugation time should be extended to ensure complete removal of cellular debris.
3. Prepare 2.5 mL ANTI-FLAG® M2 Affinity Gel (Sigma, A2220) in 20 mL empty gravity chromatography column. Allow storage buffer to drain and wash the gel according to the protocol below. After every wash allow buffer to drain completely before proceeding to the next step.

a) Wash gel with three volumes (7.5 mL) TBS.
b) Wash column three times with 2 volumes (5 mL) 0.1 M glycine-HCl pH = 3.5. Do not leave resins in glycine buffer for longer than 20 min.
c) Wash gel five times with 2 volumes (5 mL) TBS.
d) Wash gel three times with 2 volumes (5 mL) lysis buffer.

4. Resuspend resins in 10 mL of lysis buffer and transfer to the bottle containing protein supernatant. Incubate at 4°C for 30 min with gentle rotation or shaking.
5. Transfer supernatant with resins back to gravity chromatography column and allow it to pass through the column. Collect flow through. Save 5 µL for SDS-PAGE gel.
6. Wash resins three times with 10 volumes (25 mL) of TBS. Save 5 µL of each wash fraction for SDS-PAGE gel.
7. Elute bound mTet1CD protein. Resuspend beads with 1 volume (2.5 mL) of elution buffer and incubate for 5 min. Then allow buffer to flow though the column and collect the fraction. Repeat 5 times. Save 5 µL of each fraction for SDS-PAGE gel.
8. Concentrate eluted protein on Amicon Ultra 30 kD filter until concentration is around 2.5 – 5 mg/mL.
9. Buffer-exchange Bio-Spin® P-30 Gel Column (Bio-Rad, 7326231) with storage buffer and purify concentrated protein fraction. Do not apply more than 70 µL of protein solution to the column. Save 2-5 µL of protein for SDS PAGE gel.
   Note: This step is performed to remove residual Flag-peptide and protein buffer into final storage buffer.
10. Add glycerol to the concentration 30%, measure final protein concentration by Nanodrop (Thermo Fisher), aliquot and freeze in liquid nitrogen. Store protein at -80˚C. Avoid multiple freeze-thaw cycles.
11. Run collected fractions on SDS-PAGE gel to check the quality and purity of mTet1CD protein. mTet1CD MW ~ 75 kD.
    General notes: Keep all protein-containing fractions on ice during whole purification procedure. If possible perform purification at 4˚C. Try to finish the purification procedure within one day.

Supplementary Note 6: Validation of TAPS conversion with TaqαI assay

Introduction

5mC conversion after TAPS can be tested by PCR amplification of a target region which contains the TaqαI restriction site (TCGA) and subsequent TaqαI digestion. For example, 5mC conversion in our TAPS libraries can be tested based on a 194 bp amplicon containing a single TaqαI restriction site that is amplified from CpG methylated lambda DNA spike-in control. The PCR product amplified from the 194 bp amplicon is digested with TaqαI restriction enzyme and the digestion product is visualised on a 2% agarose gel. PCR product amplified on unconverted control DNA is digested by TaqαI and shows two bands on the gel. In a TAPS-converted sample, the restriction site is lost due to the C-to-T transition and so the 194 bp amplicon would remain intact. The overall conversion level can be assessed based on the ratio of the undigested and digested gel band intensities and, for successful TAPS, the conversion rate should be higher than 95%.

Reagents

- Taq DNA Polymerase with Standard Taq Buffer (NEB, M0273S)
- dNTP Set, PCR Grade, 4 x 100 µl (Qiagen, cat. no. 201912)
- UltraPure™ DNase/RNase-Free Distilled Water (Invitrogen, cat. no. 10977035)
- 194mer_Test_Forward primer: 5'-GCTGGGGAACTACAGGCT-3' (IDT)
- 194mer_Test_Reverse primer: 5'-AGAACCAGAACTCAAACTGTAC-3' (IDT)
- TaqαI (NEB, cat. no. R0149S)

Procedure

1. Set up PCR reaction according to the table below. Remember to include the unconverted control sample and non-template control.

Reagent	Volume	Final amount
DNA template	Various	2 ng
10X Standard Taq Buffer	2.5 µL	1X
dNTP [10 mM each]	0.5 µL	0.2 mM
5mC_test_Forward primer [10 µM]	0.5 µL	0.2 µM
5mC_test_Reverse primer [10 µM]	0.5 µL	0.2 µM
Taq DNA Polymerase	0.125 µL	0.625 U
Nuclease-free water	Various	NA
Total	25 µL	

Cycling conditions:

Temperature [˚C]	Time	Cycles
95	30 sec	1
95	30 sec	35
48	30 sec	
68	1 min	
68	5 min	1
4	Hold	NA

2. Check PCR products on agarose gel. A single band at 194 bp should be visible on the gel.
3. Set up the TaqαI restriction digestion:

Reagent	Volume	Final amount
194 bp PCR product	5 µL	Various
10X CutSmart buffer	1 µL	1X
TaqαI	0.2 µL	4 U
Nuclease-free water	3.8 µL	NA
Total	10 µL	

4. Incubate reactions at 65°C for 30 min.
5. Run digestion products on an 2% agarose gel. The PCR product amplified on the unconverted sample should be completely digested to yield two bands: 101 bp and 93 bp. The PCR products amplified on the TAPS-converted samples should remain undigested. Successful TAPS should have less than 5% digestion.

194 bp control sequence (5' to 3'):

GCTGGGGAACTACAGGCTGACAGTCCGGGCGGTAAATGCGTGGGGGCAGCAGGGCGATCCGGCGTCGGT
ATCGTTCCGGATTGCCGCACCGGCAGCACCG**TCGA**GGATTGAGCTGACGCCGGGCTATTTTCAGATAACCG
CCACGCCGCATCTTGCCGTTTATGACCCGACGGTACAGTTTGAGTTCTGGTTCT

Note: For designing PCR primers of TAPS-converted DNA, follow the general PCR primer designing rules and avoid CpG site at 3' end of primers. Unlike bisulfite PCR, no "C-to-T conversion" of the sequence is needed for primer design.

Supplementary Note 7: TET-Assisted Pyridine borane Sequencing (TAPS)

Introduction

TET-Assisted Pyridine borane Sequencing (TAPS) is a new method for base-resolution detection of 5mC and 5hmC. In this method, 5mC and 5hmC are first oxidized to 5caC by mTet1CD enzyme and then converted to DHU during a Pyridine borane reaction. DHU is read as T in the following PCR reaction, which enables base-resolution detection of modified cytosines after sequencing. TAPS conditions are milder compared to bisulfite sequencing and do not cause notable DNA damage.

Detailed protocol for TAPS optimized for 100 ng genomic DNA is described below.

Reagents

- EB buffer (Qiagen, 19086)
- Agencourt AMPure XP (Beckman Coulter, A63881)
- UltraPure™ DNase/RNase-Free Distilled Water (Invitrogen, 10977035)
- Ethanol (Fisher Scientific, E/0650DF/17)
- KAPA Hyper Prep Kit (KAPA Biosystems, KK8504)
- Illumina multiplexing adapters, 15 μM stock (IDT, sequences listed below)
- Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 09719-50G)
- L-Ascorbic acid (Sigma-Aldrich, 95210-50G)
- α-Ketoglutaric acid disodium salt hydrate (Sigma-Aldrich, K3752-5G)
- Dithiothreitol DTT (Fluorochem, M02712)
- NaCl (Invitrogen, AM9759)
- HEPES (Sigma, H3375)
- Adenosine 5′-triphosphate disodium salt hydrate (Sigma, A6419)
- Proteinase K (NEB, P8107S)
- Pyridine borane (Alfa Aesar, L13178.09, 95%, ~10 M)
- Sodium hydroxide (Sigma, 71960)
- Acetic acid (Sigma, A6283)
- Zymo-Spin IC (Zymo, C1004-50)
- KAPA HiFi HotStart Uracil+ ReadyMix PCR Kit (KAPA Biosystems, KK2801)
- NEBNext® Multiplex Oligos for Illumina (e.g. Index Set 1, NEB, E7335S)
- Qubit™ dsDNA HS Assay Kit (ThermoFisher, Q32854)

Preparation of DNA samples

Mix the genomic DNA of interest with the necessary control spike-ins (e.g. 0.5% methylated lambda DNA), sonicate in EB buffer and purify e.g. by AMPure XP beads to select for the desired DNA fragment size. We fragmented our DNA samples to around 200 bp using the Covaris M220 instrument and selected for 200 bp – 400 bp fragments using 0.55X – 1X AMPure XP beads.

Preparation of sequencing library with KAPA Hyper kit

Use 100 ng of fragmented and size selected DNA and prepare the end-repair and A-tailing reaction and adaptor ligation according to the manufacturer's protocol with a few modifications:

1. Use the following Illumina multiplexing adapters without the phosphorothioate bonds modifications instead of the standard KAPA index adapters for the ligation step:
   5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCT
   5'-/5Phos/GATCGGAAGAGCACACGTCT
   Both adapter oligos were obtained from IDT with HPLC purification. Anneal both adapter oligos (final 15 μM) in 10 mM Tris-Cl (pH = 8.0), 0.1 mM EDTA (pH = 8.0), 50 mM NaCl with the following program: 2 min at 95°C, 140 cycles of 20 sec at 95°C (decrease temperature 0.5°C every cycle) and hold at 4°C. Annealed 15 μM Illumina multiplexing adapters should be aliquoted into small single-use vials and stored at -80°C.
   Note: Methylated adapters for bisulfite sequencing cannot be used for TAPS and other versions of Illumina adapters might affect the performance of TAPS.
2. Extend ligation step to 30 min.
3. After post-ligation clean-up, elute the adapter-ligated DNA sample in 20 μL nuclease-free water.

mTet1 oxidation

1. Prepare mTet1CD oxidation buffer: 167 mM HEPES pH = 8.0, 333 mM NaCl, 3.3 mM α-KG, 6.67 mM L-AA, 4 mM ATP, 8.33 mM DTT. After preparation, the oxidation buffer should be aliquoted and stored at -80°C. Buffer aliquots are single-use reagents and should not be freeze-thawed multiple times.
2. Set up the oxidation reaction as follows (on ice):

Reagent	Volume	Final amount
Adapter-ligated DNA	20 μL	Up to 100 ng

mTet1CD oxidation buffer	15 µL	1X
1.5 mM Fe	3.33 µL	100 µM
mTet1CD	Various	4 µM
Nuclease-free water	Various	NA
Total	50 µL	

3. Incubate reaction at 37°C for 80 min.
4. Add 1 µL Proteinase K (0.8 U) to the oxidation reaction and incubate for 1 h at 50°C.
5. Purify reaction on Bio-Spin® P-30 Gel Column and then on 1.8X Ampure XP beads according to manufacturer's protocol. Elute in 35 µL nuclease-free water.
6. For more complete oxidation, repeat the whole procedure for a second round.
   Note: Typical DNA recovery after mTet1 oxidation is ~75%.

Pyridine borane reaction

1. Prepare 3 M sodium acetate buffer solution pH = 4.3 by adjusting pH of 4 M acetic acid with 5 M sodium hydroxide solution. Add water if necessary to obtain the final concentration 3 M.
2. Add 10 µL of 3 M sodium acetate solution pH = 4.3 and 5 µL of 10 M pyridine borane to 35 µL DNA sample. Incubate at 37°C and 850 rpm in a ThermoMixer (Eppendorf) for 16 hours.
   Note: Final reaction contains 600 mM NaAc pH = 4.3 and 1 M pyridine borane.
3. Purify the reaction product on Zymo-IC column with PB buffer according to the manufacturer's protocol.
   Note: Expected DNA recovery after this step is 50-70%.

Library amplification

1. Amplify the sequencing library with KAPA HiFi HotStart Uracil+ ReadyMix PCR Kit for 5 cycles according to the manufacturer's protocol. Use index primers in NEBNext® Multiplex Oligos for Illumina kit instead of the Library Amplification Primer Mix.
2. Purify with 1X Ampure XP beads according to the manufacturer's protocol. Check the quality and purity of the final library e.g. on 2% agarose gel.
   Note: If starting with 100 ng of DNA for library construction, up to 5 cycles of PCR can be performed to generate a final library of > 30 nM.