Supporting Information

Bisulfite-free, single-base resolution analysis of 5-hydroxymethylcytosine in genomic DNA by chemical-mediated mismatch

Yafen Wang[‡,†], Xiong Zhang[‡,†], Fan Wu[†], Zonggui Chen[#] & Xiang Zhou[*,†]

[†]College of Chemistry and Molecular Sciences, Key Laboratory of Biomedical Polymers of Ministry of Education, The Institute for Advanced Studies, Hubei Province Key Laboratory of Allergy and Immunology, Wuhan University, Wuhan, Hubei, 430072, P. R. China

[#]College of Life Science, Wuhan University, Wuhan, Hubei, 430072, P. R. China

[*]Corresponding Author, E-mail: xzhou@whu.edu.cn; Fax: +86-27-68756663; Tel: +86-27-68756663

[‡]These authors contributed equally.

List of Contents:

1. Materials and methods

Materials and chemicals. All chemicals were purchased from Beijing Innochem Sci. & Tech. Co. Ltd. (Beijing, China) unless mentioned otherwise. The shorter oligonucleotides containing 5-formylcytosine (ODN-5fC), dNTP (N = A, T, C, G), Hot Start Taq polymerase and MightyAmp DNA polymerase were bought from Takara Biotechnology Co., Ltd. (Dalian, China). 2× Hieff™ PCR SYBR® Green Master Mix were ordered from YEASEN (Shanghai, China). Dream Taq polymerase was obtained from ThermoFisher Scientific (USA). dhmCTP and dfCTP were purchased from TriLink Biotechnologies (USA). DNA Clean & Concentrator™-5 kit and Degradase Plus were bought from Zymo Research (USA). Micro Bio-Spin™ P-6 Gel Columns were purchased from Bio-Rad (USA). pClone007 Simple Vector Kit was ordered from TSINGKE Biological Technology (Beijing, China). The mESCs culture medium was bought from Cyagen Biosciences (China). The nucleic acid stains Super GelRed (NO.: S-2001) was bought from US Everbright Inc. (Suzhou, China). All of the unmodified oligonucleotides were synthesized and purified by GeneCreate Co., Ltd. (Wuhan, China). VAHTS DNA Clean Beads were purchased from Vazyme (Nanjing, China). NEBNext® Ultra™ II DNA Library Prep Kit for Illumina was brought from New England BioLabs (USA). Qubit™ dsDNA HS Assay Kit was purchase from ThermoFisher Scientific (USA). The NGS data were obtained from Whbioacme Co. Ltd and GENEWIZ.

General oxidation of synthetic DNA. The mixture of DNA samples of 80bp ds-X (X= hmC and fC) after being ligated with adapter were subjected to oxidation. Generally, 1.25 μL freshly prepared NaOH (1 M) was added into the mixture for the final volume was 24 μL. The mixture was denatured at 37°C for 0.5 h. After denaturing, 1 μL $KRuO_4$ solution (15 mM in 0.05 M NaOH) was added and the reaction was processed on ice for 1 hour, with occasional vortexing. Then the reaction was purified by Micro Bio-Spin™ P-6 Gel Columns in SSC Buffer pre-washed with ddH_2O (3 × 500 μL) for removing NaOH and $KRuO_4$.

General oxidation of genomic DNA. As for oxidation of genomic DNA, genomic DNA (1 μg or less) was first denatured by heating at 95°C for 5 min and then chilling on ice for 2 min. After that, NaOH (final concentration was 50 mM) was added into the genomic DNA for 30 min at 37°C for secondary denaturing. Then 1 μL $KRuO_4$ solution (15 mM in 0.05 M NaOH) was added and the reaction was processed on ice for 1 hour, with occasional vortexing. Then the reaction was purified by Micro Bio-Spin™ P-6 Gel Columns in SSC Buffer pre-washed with ddH_2O (3 × 500 μL) for removing NaOH and $KRuO_4$.

Chemical labeling reaction. For the reaction between ODNs and azi-BP, 1 μL ODN (100 μM) was added into the mixture of 5 μL HEPES buffer (1 M, pH 7.4) and 19 μL ddH_2O and 25 μL azi-BP (100 mM, dissolved in DMSO). After vortexing and centrifugation, the mixture was incubated in a thermos-shaker (Ningbo Biocotek Scientific Instrument Co., Ltd., China, 1500 r.p.m.) at 56℃ for 18 h. The product was purified by DNA Clean & Concentrator™-5 kit. The purified product was incubated with 10 μL DBCO-PEG4-biotin (10 mM) or DBCO-S-S-PEG3-biotin (20 mM) at 37℃ for 2 h.

Adapter ligation of DNA. 80 bp ds-X (X = C, 5mC, 5hmC and 5fC, E-DNA-C, E-DNA-hmC, E-

DNA-fC each 10 ng) and calf thymus DNA (10 μg) were mixed and subjected to adapter ligation using the standard protocol according to the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina (New England Biolabs), except the use of the modified adapter solution (2.5 μL) instead of the adapter in the kit. Then the mixture was purified by VAHTS DNA Clean Beads for removing the excess adapter.

Enrichment and qPCR analysis of 5hmC in the mixture. The adapter ligated DNA sample was oxidized according to the protocol described above (see General oxidation of synthetic DNA). Then the oxidized system was labeled with azi-BP and DBCO-S-S-PEG3-biotin for enrichment.

VAHTS DNA Clean Beads for purification. The beads should warm to room temperature for at least 0.5 h before use. Add 0.9X (volume) of resuspended beads to the reaction and mix thorough by pipetting up and down. Then the mixture was incubated at room temperature for 10 minutes. Following that, the tube was placed on magnetic stand to removing the supernatant. The beads were washed by 80% freshly prepared ethanol (2 × 200 μL) and then the beads were dried in air. Lastly, the targeted DNA was eluted by 0.1 × TE.

Illumina sequencing for the model DNA of 80 bp ds-5hmC. 80 bp ds-5hmC was ligated with the adapter as described above. Then the DNA was oxidized using the above protocol. After that, the DNA was incubated with azi-BP and DBCO-S-S-PEG3-biotin. The purified DNA was subjected to NEB Illumina library preparation using NEBNext® Ultra™ II DNA Library Prep Kit (NEB, USA) and the library was used for next-generation sequencing.

The protocol of qPCR analysis of 5fC. qPCR was performed using a CFX-96 Real-Time System (Bio-Rad, USA). The mixture contained 1 μL template of the labeled DNA or unlabeled DAN, 1 μL forward primer (10 μM), 1 μL reverse primer (10 μM), 10 μL 2 × Hieff™ PCR SYBR® Green Master Mix and 7 μL ddH_2O to give a final volume of 20 μL. The mixture was subject to qPCR according to the following thermal cycle: 95 °C for 5 min, 40 cycles of (95 °C for 10 s, 60 °C for 30 s, 72 °C for 20 s).

Chemical labeling-mediated qPCR assay. Labeled DNA was used as template for qPCR assay. The DNA without treatment were as control samples. The Ct values corresponding to labeled DNA and unlabeled DNA are named as Ct_1 and Ct_0, respectively. $\Delta Ct = Ct_1 - Ct_0$.

DNA enrichment protocol. The enrichment procedure was performed according to the manufacturer's instructions with minor modifications to the use of Dynabeads™ MyOne™ Streptavidin C1 (ThermoFisher Scientific). Briefly, 1× binding and washing (B&W) buffer (pH 7.5) was added with 0.05% Tween-20. Beads (20 μL) were washed with 500 μL 1 × B&W buffer for three times and then resuspended in 20 μL 2 × B&W buffer. Input DNA (the mixture of E-DNA-C, E-DNA-hmC and E-DNA-fC) and calf thymus (10 μg, sonicated into 200-400 bp) were mixed and added ddH_2O to a final volume of 20 μL and then mixed with magnetic beads. Then pipette the entire system up and down gently. The mixture was incubated at room temperature for 30 minutes on a nutator. After that, the tube was placed on a magnet for 5 minutes before discarding the supernatant, and beads were washed with 500 μL 1 × B&W buffer for five times.

Finally, the biotinylated DNA was released by DTT (50 mM) cleaving disulfide linkage and purified with DNA Clean & Concentrator™-5 kit.

qPCR analysis for chemical enrichment studies. qPCR was performed using a CFX-96 Real-Time System (Bio-Rad, USA). The system was composed of 2 × Hieff™ PCR SYBR® Green Master Mix (10 μL), forward primer (1 μL, 10 μM), reverse primer (1 μL, 10 μM), template and ddH_2O. The mixture was performed to qPCR according to the following thermal cycle: initial denaturation at 95°C for 5 min, 40 cycles of (95 °C for 10 s, 60 °C for 30 s, 72 °C for 20 s). The concentration of DNA was quantified by calculation with calibration lines of known concentration of input ODNs.

Mouse embryonic stem cells culture and DNA extraction. Mouse embryonic stem cells (mESCs) were cultured in Mouse Embryonic Stem Cell (mESC) Basal Medium with 15% mESC-qualified fetal bovine serum (FBS), penicillin/streptomycin, glutamine, nonessential amino acid, LIF and 2-mercaptoethanol were plated in culture dishes pretreated with 0.1% gelatin and mouse embryonic fibroblast then incubated in humidified 37°C incubator supplied with 5% CO_2. Genomic DNA was extracted from the harvested cells using a DNeasy® Blood & Tissue Kit (Qiagen, Germany) following the manufacturer's instructions.

2. Sanger sequencing analysis of multiple 5hmC sites in model DNA.

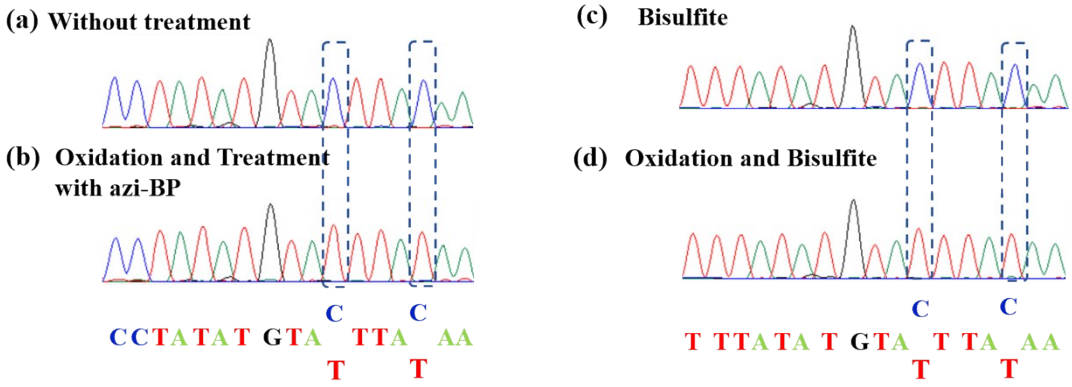
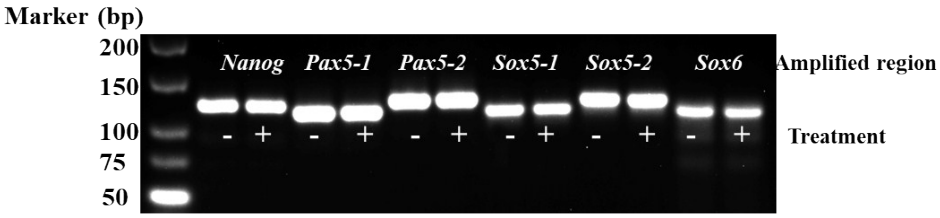


Figure S1. Model ODN (seq-2 ds-ODN-5hmC) to verify the feasibility of the method CAM-seq. (a) The Sanger sequence results of original sequence. (b) The Sanger sequence result of model ODN by CAM-seq. (c) The Sanger sequence results of bisulfite. (d) The Sanger sequence results of oxidation and bisulfite.

3. Agarose gel electrophoresis analysis of genomic DNA PCR product



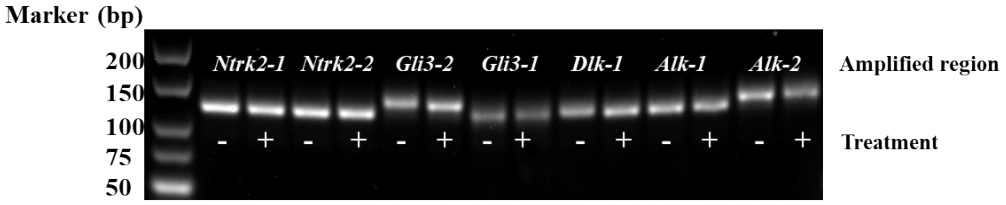


Figure S2. Agarose gel electrophoresis analysis of genomic DNA PCR product from mouse ES cells. "-" represent the genomic DNA without treatment, "+" represent the genomic DNA was oxidized by $KRuO_4$ and then treated with azi-BP.

4. Sanger sequencing analysis of 5hmC in genomic DNA of mESCs

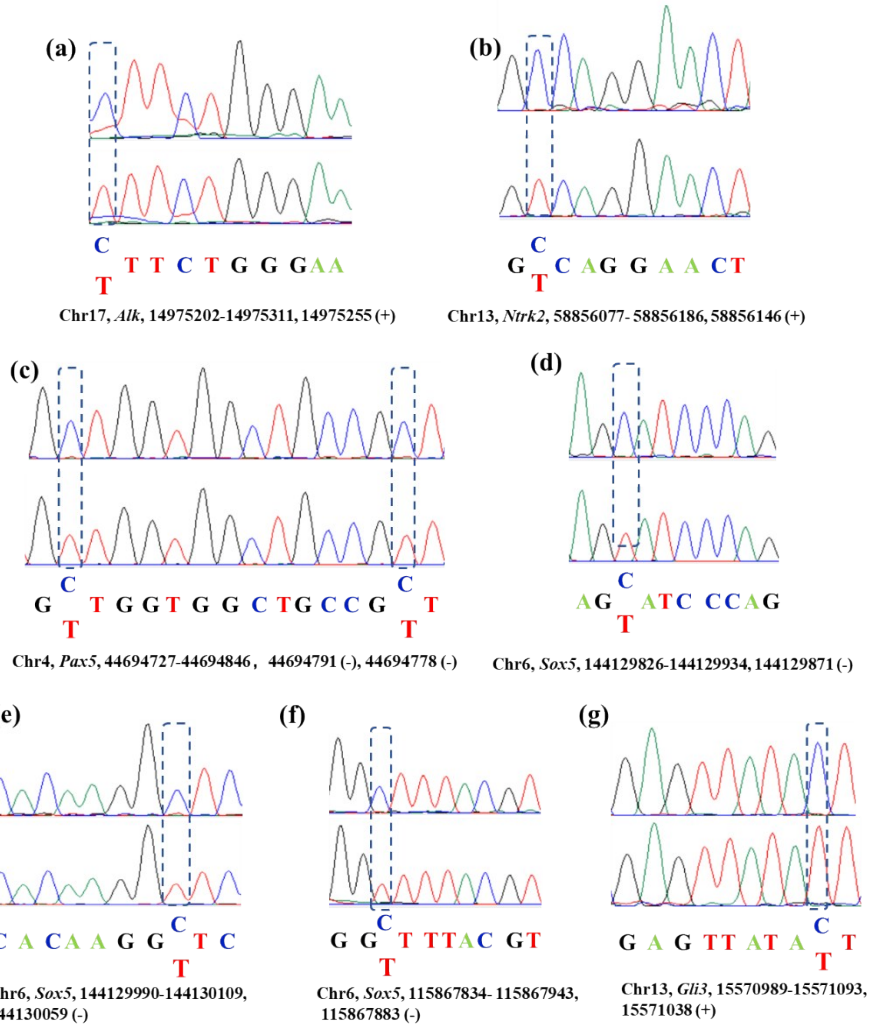


Figure S3. Sanger sequencing analysis 5hmC by CAM-seq in mESCs. The base positions originally from 5fC is surrounded by blue dotted lines.

5. Calibration line of qPCR for enrichment analysis

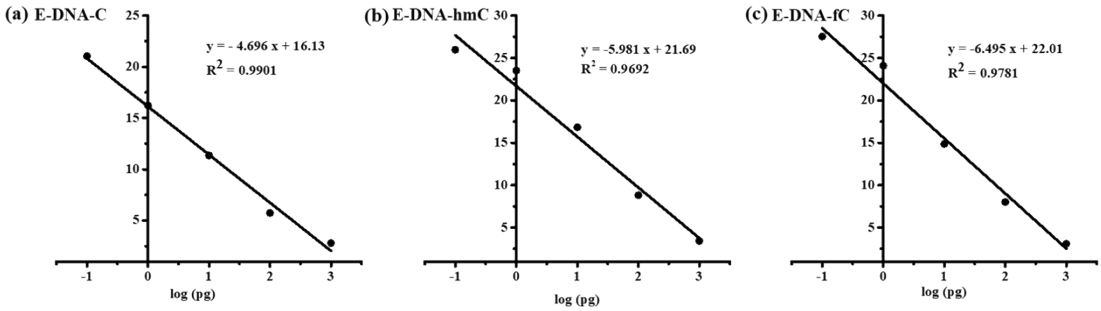


Figure S4. Example calibration line of E-DNA-C, E-DNA-hmC and E-DNA-fC for enrichment analysis.

6. Ct value of different DNA polymerase and different labeling compounds

We tested the Ct values for different DNA polymerase utilizing a double strand DNA containing eight 5hmC sites. Apart from HieffTM DNA Polymerase (YEASEN), SYBR® Premix Ex TaqTM (Tli RNaseH Plus) (Takara), MightyAmp DNA Polymerase (Takara), Taq DNA Polymerase (recombinant) (Thermo Fisher), Dream Taq DNA Polymerase and NEBNext Ultra II Q5 Master Mix were also used for the Ct value test of the labelled ODN containing 5hmC. As shown in Figure S5a and Figure S5b, the results indicated that when 5fC was labelled with azi-BP and further labelled with DBCO-PEG4-biotin, the added group could hamper the reactivity of DNA polymerase (SYBR® Premix Ex TaqTM and MightyAmp DNA Polymerase). However, when using the Taq DNA Polymerase (recombinant), Dream Taq DNA Polymerase and NEBNext Ultra II Q5 Master Mix, no fluorescence was found in qPCR (Figure S5c). We speculated that the presence of SYBR Green Ⅰ might inhibit the activity of certain enzymes under certain circumstances. To further verify the conjecture, we tested the product of polymerase chain reaction (PCR) in the presence and absence of SYBR Green Ⅰ by 3% agarose gel electrophoresis analysis (Figure S6). We found the product of 80 bp dsDNA in the absence of SYBR Green Ⅰ, however, no dsDNA was found in the presence of SYBR Green Ⅰ. Therefore, SYBR Green Ⅰ selectively inhibit the activity of certain enzymes, such as Taq DNA Polymerase (recombinant), Dream Taq DNA Polymerase and NEBNext Ultra II Q5 Master Mix. 5fC represents the oxidized 5hmC (5fC) and 5fC + D-PEG-P represents oxidized 5hmC (5fC) was labelled with azi-BP and further incubated with DBCO-PEG4-biotin.

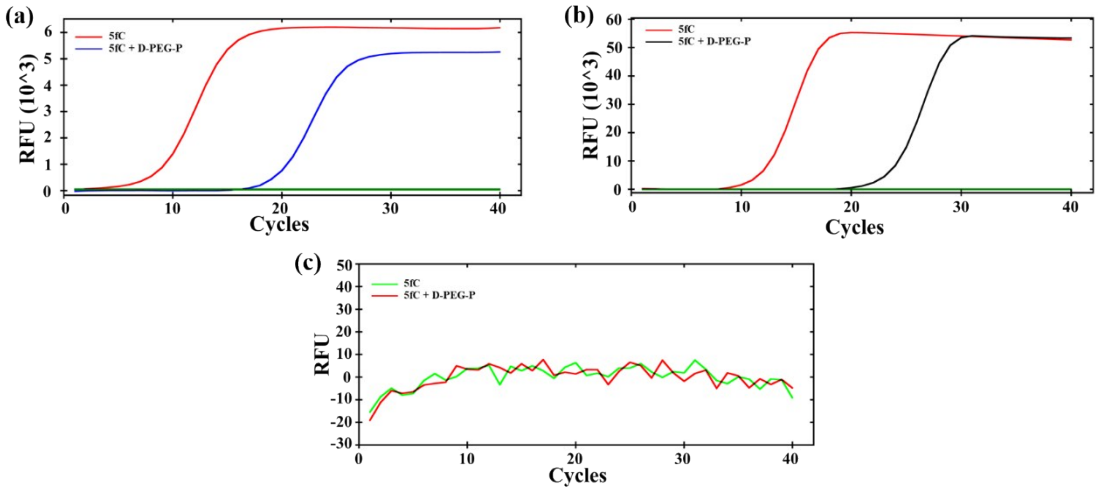


Figure S5. The Ct values for different DNA polymerase. (a) SYBR® Premix Ex Taq™ (Takara), (b) MightyAmp DNA Polymerase (Takara), (c) NEBNext Ultra II Q5 Master Mix.

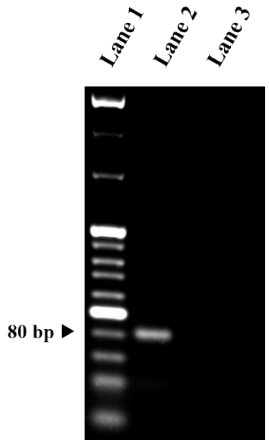


Figure S6. Agarose gel electrophoresis analysis of the product of polymerase chain reaction (PCR) in the presence and absence of SYBR Green Ⅰ. Lane 1: 20bp DNA Ladder, Lane 2: in the absence of SYBR Green Ⅰ, Lane 3: in the presence of SYBR Green Ⅰ.

The Ct values of different labeling compounds were tested. Similarly, a double strand DNA containing eight 5hmC sites was also chosen for the experiment. The results showed (Figure S7) that when 5fC was labeled with azi-BP or further labeled with DBCO-PEG4-biotin, DBCO-S-S-PEG3-biotin, as well as subsequently treated with DTT, all the added group could hamper the reactivity of DNA polymerase under the treatment of MightyAmp DNA Polymerase. **5fC** represents the oxidized 5hmC (5fC), **5fC + azi-BP** represents oxidized 5hmC (5fC) was incubated with azi-BP, **5fC + D-S-S-P** represents oxidized 5hmC (5fC) was labelled with azi-BP and further incubated with DBCO-S-S-PEG3-biotin, **5fC + D-S-S-P + DTT** represents **5fC + D-S-S-P** was treated with DTT, **5fC + D-PEG-P** represents oxidized 5hmC (5fC) was labelled with azi-BP and further incubated with DBCO-PEG4-biotin.

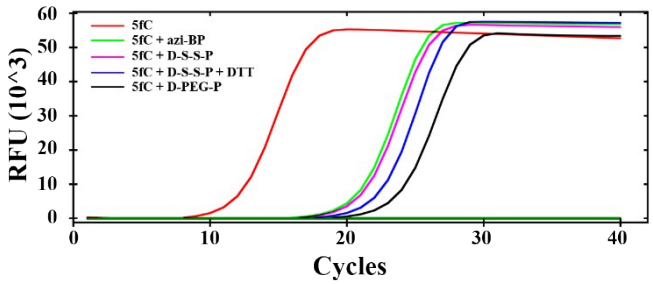


Figure S7. The Ct values of different labeling compounds under the same condition in the presence of MightyAmp DNA Polymerase.

7. Different modification level of 5hmC in model sequences for NGS

5hmC model sequences containing different modification level was tested by NGS. The procedure was as follows. (a) Single specific site of 5hmC in DNA (Ratio-seq-C, Ratio-seq-5hmC) with different modification ratio (0, 25, 50, 75, 100% 5hmC level at a specific site) was ligated to the modified NGS adapter and then oxidized by $KRuO_4$ and purified with Micro Bio-Spin™ P-6 Gel Columns (Bio-Rad). (2) Oxidized 5hmC was labeled with azi-BP and purified with DNA Clean & Concentrator™-5 kit (Zymo Research). (3) The purified DNA annealed in PCR system for library preparation and NEBNext Ultra II Q5 Master Mix was placed with MightyAmp polymerase. The results indicated that with the increase of modification level of 5hmC, the C-to-T conversion ratio increased. However, without enrichment, the C-to-T conversion ratio for 5hmC is only ~ 33% (Figure S8e), the low conversion ratio may be caused by the incompletely label of 5hmC. So to improve the C-to-T conversion ratio, the step of enrichment is important.

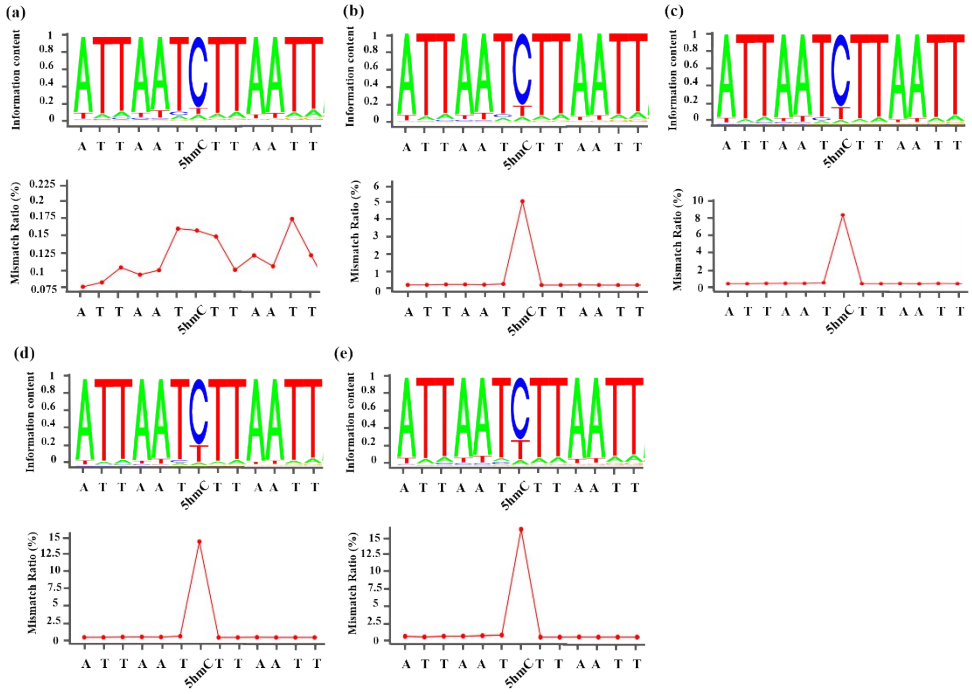


Figure S8. Oxidized-chemical-labeling-induced C-to-T conversion of 5hmC in a double-stranded 5hmC-containing different modification level. (a)-(e) respresent 0, 25, 50, 75, 100% 5hmC level at

a specific site respectively. Top: Sequence logo, bottom: Mismatch ratio of CAM-Seq.

8. The oxidation efficiency and degradation of $KRuO_4$

It is reported that $KRuO_4$ oxidation causes significant DNA degradation[1]. A 15 mer ODN containing 5hmC (Phos-GACTCAA5hmCAGCCGTA-phos, named ODN-5hmC-phos) was chosen to test the oxidation efficiency. ODN-5hmC was oxidized by $KRuO_4$ as the protocol reported previously[1, 2]. Initial measurements of overall DNA yield by HPLC and PAGE（Figure S9）indicated that almost no DNA degradation was found during oxidation. Then the oxidized ODN-5hmC-phos was digested for HPLC-MS analysis. After oxidation, the peak of 5hmC was not found and an obvious peak of 5fC appeared (Figure S10), indicating the high oxidation efficiency up to about 100%. Subsequently, we chose another 15 mer ODN containing 5hmC without phosphate group protection (5'-GACTCAA5hmCAGCCGTA-3', named ODN-5hmC) to be oxidized by $KRuO_4$ in the same way. Instead, we observed about 80% DNA was degraded by combining HPLC with PAGE analysis (Figure S11). We speculated that the 3- and 5-termial with phosphate groups could improve the stability of DNA and protect DNA from degradation to some extent.

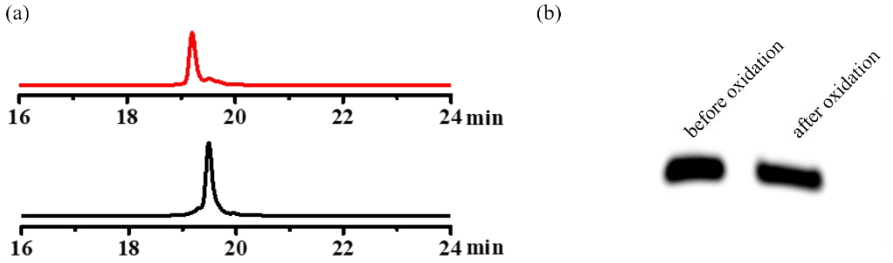
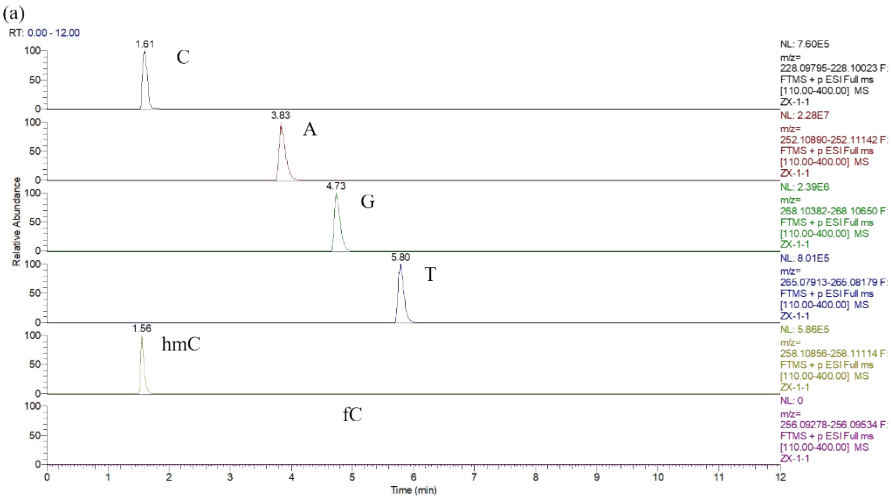


Figure S9. The strategy for testing ODN-5hmC degradation efficiency oxidized by $KRuO_4$. (a) RP-HPLC traces of ODN-5hmC before (red) and after (black) oxidation by $KRuO_4$. (b) Denaturing PAGE analysis of ODN-5hmC before and after oxidation by $KRuO_4$ with the same quantity.



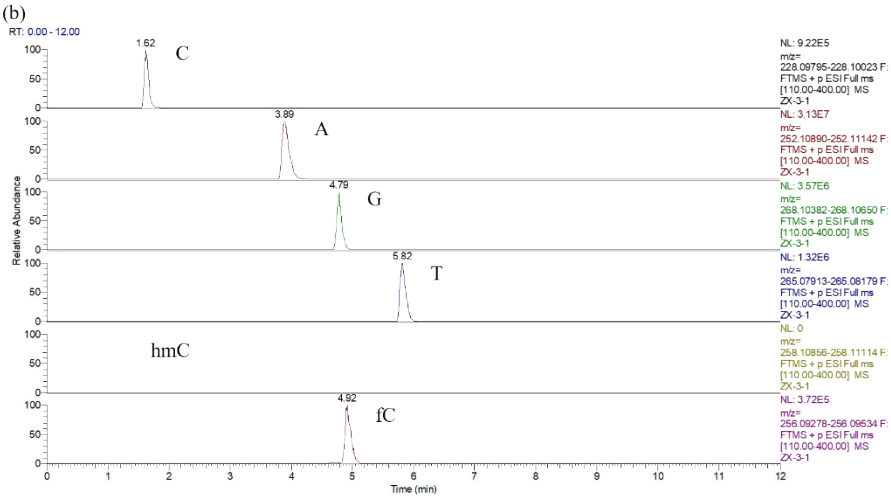


Figure S10. HPLC-MS extracted $[M+H]^+$ ion count for C, A, G, T, 5-hmC and 5-fC deoxynucleosides after digestion of a) ODN-5hmC, b) ODN-5hmC after treatment with KRuO4.

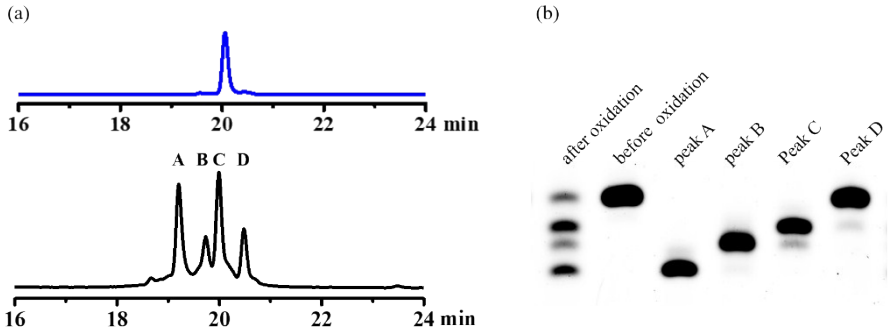


Figure S11. The strategy for testing 5hmC-ODN degradation efficiency oxidized by $KRuO_4$. (a) RP-HPLC traces of 5hmC-ODN before (blue) and after (black) oxidation by $KRuO_4$. (b) Denaturing PAGE analysis of 5hmC-ODN before and after oxidation by $KRuO_4$.

9. The sequences used in this study

Table S1. The sequences of the oligodeoxynucleotides used in this study.

Note: [a] The corresponding templates were used listed in Table S2. [b] The corresponding primers were used listed shown in Table S3, F is short for forward, R is short for reverse.

ODNs	Sequence
ODN-5fC	AGAT5fCGTAT
ODN-5hmC-Phos	Phos-GACTCAA5hmCAGCCGTA-phos
ODN-5hmC	GACTCAA5hmCAGCCGTA
E-DNA-C (Template 1)[a] F 1, R 1)[b]	5'-GGGTTTTATTATTTTAATTAATATTATATTATGTGTATAACGTGTATTGTGTGTTATAATATTGAGGGAGAAGTGGTGAG-3' Complementary sequence 3'-CCCAAAATAATAAAATTAATTATAATATAATACACATATTGCACATAACACACAATATTATAACTCCCTCTTCACCACTC-5'
E-DNA-hmC	5'-

(Template 2)[a] F 2, R 2)[b]	TCCTCCTACATCATTCCTCTCTAACCCCTTATATGTA5hmCTTAG AAT5hmCAATTGAGTGATTGAAGGTAGTTAGTGGTGGTAGA-3' Complementary sequence 3'- AGGAGGATGTAGTAAGGAGAGATTGGGGAATATA5hmCATGAA T5hmCTTAGTTAACTCACTAACTTCCATCAATCACCACCATCT-5'
E-DNA-fC (Template 3)[a] F 3, R 3)[b]	5'- TTCCTTCACCATCTCAACCAATATTATATTAAATTATAAT5fCTG TATT5fCTGTTTTATAATATTGAGGGAGAAGAATGGTGGA-3' Complementary sequence 3'- AAGGAAGTGGTAGAGTTGGTTATAATATAATTTAATATTAGA5f CATAAGA5fCAAAATATTATAACTCCCTCTTCTTACCACCT-5'
seq-1 ds-ODN-5hmC (Template 4)[a] F 4, R 4)[b]	5'- TTCTTCTACATCATCTCCCTCTAACCTCCTATATGTA5hmCTTAG AATTAATTAGGTAGATTAGAGAGTAGTAGTGGTGAGGA-3' Complementary sequence 3'- AAGAAGATGTAGTAGAGGGAGATTGGAGGATATACATGAATC TTAATTAATCCATCTAATCTCTCATCATCACCACTCCT-5'
seq-2 ds-ODN-5hmC (Template 5)[a] F 4, R 4)[b]	5'- TTCTTCTACATCATCTCCCTCTAACCTCCTATATGTA5hmCTTA 5hmCAATTAATTAGGTAGATTAGAGAGTAGTAGTGGTGAGGA-3' Complementary sequence 3'- AAGAAGATGTAGTAGAGGGAGATTGGAGGATATACATGAATG TTAATTAATCCATCTAATCTCTCATCATCACCACTCCT-5'
seq-1 ds-ODN-C (Template 4)[a] F 4, R 4)[b]	5'- TTCTTCTACATCATCTCCCTCTAACCTCCTATATGTACTTAGAAT TAATTAGGTAGATTAGAGAGTAGTAGTGGTGAGGA-3' Complementary sequence 3'- AAGAAGATGTAGTAGAGGGAGATTGGAGGATATACATGAATC TTAATTAATCCATCTAATCTCTCATCATCACCACTCCT-5'
DNA-fC (Template 2)[a] F 2, R 2)[b]	5'- TCCTCCTACATCATTCCTCTCTAACCCCTTATATGTA5fCTTAGA AT5fCAATTGAGTGATTGAAGGTAGTTAGTGGTGGTAGA-3' Complementary sequence 3'- AGGAGGATGTAGTAAGGAGAGATTGGGGAATATA5fCATGAAT 5fCTTAGTTAACTCACTAACTTCCATCAATCACCACCATCT-5'
dsODN-hmC-8	5'- TTCTTCTACATCATCTCCCTCTAACCTCCTAT5hmCGTA5hmCGT TA5hmCGATA5hmCGTTAGGTAGATTAGAGAGTAGTAGTGGTGA GGA -3' Complementary sequence 3'- AAGAAGATGTAGTAGAGGGAGATTGGAGGATAG5hmCATG5hm CAATG5hmCTATG5hmCAATCCATCTAATCTCTCATCATCACCAC TCCT -5'

Ratio-seq-C (Template 6)[a] F 5, R 5)[b]	5'- CTAAATCTACTAAATCCTCTAAATCTATTCTATAAATTAATCTT AATTAAAGGTAGTAGTAGTAGATATAAGATGATAGG-3' Complementary sequence 3'- GATTTAGATGATTTAGGAGATTTAGATAAGATATTTAATTAGA ATTAATTTCCATCATCATCATCTATATTCTACTATCC-5'
Ratio-seq-5hmC (Template 6)[a] F 5, R 5)[b]	5'- CTAAATCTACTAAATCCTCTAAATCTATTCTATAAATTAAT5hm CTTAATTAAAGGTAGTAGTAGTAGATATAAGATGATAGG-3' Complementary sequence 3'- GATTTAGATGATTTAGGAGATTTAGATAAGATATTTAATTAGA ATTAATTTCCATCATCATCATCTATATTCTACTATCC-5'

10. The templates used for synthesis of dsDNA

Table S2. The sequences of the templates used in this study.

ODNs	Sequence (5'-3')
Template 1	GGGTTTTATTATTTTAATTAATATTATATTATGTGTATAACGTGT ATTGTGTGTTATAATATTGAGGGAGAAGTGGTGAG
Template 2	TCCTCCTACATCATTCCTCTCTAACCCCTTATATGTACTTAGAAT CAATTGAGTGATTGAAGGTAGTTAGTGGTGGTAGA
Template 3	TTCCTTCACCATCTCAACCAATATTATATTAAATTATAATCTGTA TTCTGTTTTATAATATTGAGGGAGAAGAATGGTGGA
Template 4	TTCTTCTACATCATCTCCCTCTAACCTCCTATATGTACTTAGAATT AATTAGGTAGATTAGAGAGTAGTAGTGGTGAGGA
Template 5	TTCTTCTACATCATCTCCCTCTAACCTCCTATATGTACTTACAATT AATTAGGTAGATTAGAGAGTAGTAGTGGTGAGGA
Template 6	CTAAATCTACTAAATCCTCTAAATCTATTCTATAAATTAATCTTA ATTAAAGGTAGTAGTAGTAGATATAAGATGATAGG

11. The sequences of primers used for synthesis of dsDNA

Table S3. The sequences of the primers used in this study.

ODNs	Sequence (5'-3')
Forward primer 1	GGGTTTTATTATTTTAATTAATATTATATT
Reverse primer 1	CTCACCACTTCTCCCTCAAT
Forward primer 2	TCCTCCTACATCATTCCTCTCTAACCCCT
Reverse primer 2	TCTACCACCACTAACTACCTTCAATCACTC
Forward primer 3	TTCCTTCACCATCTCAACCAATAT
Reverse primer 3	TCCACCATTCTTCTCCCTCAATA
Forward primer 4	TTCTTCTACATCATCTCCCTCTAACCTCCT
Reverse primer 4	TCCTCACCACTACTACTCTCTAATCTACCT
Forward primer 5	CTAAATCTACTAAATCCTCTAAATCTATTC
Reverse primer 5	CCTATCATCTTATATCTACTACTACTACCT
BS-Forward primer	GGGTTTTTTTATATTATTTTTTTTTAATTTTTT

12. PCR primers that were used for the CAM-Seq of genomic DNA of mouse ES cells.

No.	Gene/ Amplified region	5hmC Position	Strand	Sequence
1	CDS_*Nanog*/ chr6: 122713337-122713453	122713406 (+)	F	ACCGCTCAGTCCTGGAA
			R	CTCCAAATCACTGGCAGA
			BS-F	GGGATTGTTTAGTTTTGGAATGG
			BS-R	GGCTCCAAATCACTAAC
2 (Sox-1)	INTRON_*Sox5*/ chr6: 144129826-144129934	144129871 (-)	F	GATCGACGGAGACTGGAA
			R	CAGGATCTCCCTGCTTCA
3 (Sox-2)	INTRON_*Sox5*/ chr6: 144129990-144130109	144130059 (-)	F	ATCGCTCTCGCTGGGGGCCA
			R	CTTGATCCCCTGCACAGACC
4 (Sox-3)	UTR-5_*Sox6*/ chr7: 115867834-115867943	115867883 (-)	F	CTGAAACCCAGTGGCATC
			R	CGCACAAAGCTGTGATTG
5	INTRON_*Pax5_2*/ chr4: 44694526-44694635	44694568 (-)	F	TGGCATGCGTCAGCCCAGGG
			R	ACGTGTCAGTCAAACGGAAG
			BS-F	ATGTGTTAGTTAAATGGAAG
			BS-R	TAACATACATCAACCCAAAACTCCC
6	INTRON_*Pax5_3*/ chr4: 44694727-44694846	44694791, 44694778 (-)	F	TCATCTCAGCAAAACAGCCG
			R	CCTCTTGACTCCCAACCTAA
7 (CDS-1)	CDS_*Dlk1*/ chr12: 109459804-109459914	109459859 (+)	F	GATTCGTCGACAAGACCTG
			R	TGCACAGACACTCGAAGC
			BS-F	GGGGATTTGTTGATAAGATTTG
			BS-R	GGGTACACAAACACTCAAAACTCAC
8 (Gli3-1)	INTRON_*Gli3*/ chr13: 15570989-15571093	15571038 (+)	F	CCAGTTGCCTCAAACTCC
			R	ATACTGCGCCCGTAATGT
9 (Gli3-3)	INTRON_*Gli3*/ chr13: 15714315-15714434	15714381 (+)	F	TAAATGGACTTGTGTCTTCA
			R	GATCGAGGAGACAGAACAAC
			BS-F	GGGTAAATGGATTTGTGTTTTTA
			BS-R	GGGAATCAAAAAAACAAAACAAC
10 (Alk-1)	INTRON_*Alk*/ chr17: 14975202-14975311	14975255 (+)	F	TGTTTGGACCTCCTCCTG
			R	GCGACATCTTCGACTCCA
11 (Ntrk-2)	INTRON_*Ntrk2*/ chr13: 58856077- 58856186	58856146 (+)	F	AGCACCGTAGCTTTTGGAAG
			R	GCCAAGGCGGGATAAACCAC

Reference

1. M. J. Booth, M. R. Branco, G. Ficz, D. Oxley, F. Krueger, W. Reik and S. Balasubramanian, *Science*, 2012, **336**, 934.
2. F. Kawasaki, S. Martínez Cuesta, D. Beraldi, A. Mahtey, R. E. Hardisty, M. Carrington and S.

Balasubramanian, *Angew. Chem. Int. Ed.*, 2018, **130**, 9842-9844.