Supplementary Information

Oxime Formation Coordination-Directed detection of genome-wide thymine oxides with nanogram-scale sample input

Feng Xiao[a,†], Qi Wang[a,†], Kaiyuan Zhang[a], Chaoxing Liu[b], Guangrong Zou[a,*] and Xiang Zhou[a,*]

[a] College of Chemistry and Molecular Sciences, Key Laboratory of Biomedical Polymers-Ministry of Education, Wuhan University, Wuhan, Hubei, 430072, P. R. China. E-mail: xzhou@whu.edu.cn

[b] University of California, Riverside Department of Chemistry

[*] To whom correspondence should be addressed. Tel: +86-27-68756663; Fax: +86-27-68756663; Email: xzhou@whu.edu.cn

[†] These authors contributed equally.

Experimental section

Materials and reagents

All chemical reagents, including organic solvent of analytical grade, materials used to synthesize the product AQA[1] and other related compounds were purchased from Sigma-Aldrich (USA). Oligos containing 5fU or 5fC were obtained from Takara Biotechnology (Dalian, China). 15 nt-AP was synthesized from 15 nt-U (one T site was replaced by uracil) treating with Uracil DNA Glycosylase (Invitrogen, USA). d5fUTP and d5hmUTP were synthesized based on the previous method.[2] All sequences of the oligos and DNA were listed in Supplementary Table S1. The electrophoretic apparatus was bought from Liuyi Instrument Factory (DYCZ- 22A, Beijing, China). Super GelRed dye was obtained from US Everbright Inc. (Suzhou, China) and SYBR Green I was bought from GOYOO BIOTECH CO.LID (Nanjing, China). MALDI-TOF-MS spectra were measured on the MALDI-TOF-MS instrument (Shimadzu, Japan). 3-hydroxypicolinic acid (3-HPA) was used as the assisted matrix. Liquid chromatography-mass spectrometry (LC-MS) data were collected with the Agilent™ 1220 Infinity LC combined with the 6120 Single Quadrupole mass spectrometer (Agilent Technologies). 80-bp dsDNAs (containing 5hmU or 5fU sites) were prepared by incorporating dhmUTP or dfUTP during the process of PCR amplification. Phi29 DNA polymerase was bought from NEB ENGLAND BioLabs (USA) and T4 DNA Ligase was obtained from Invitrogen (USA). DNA concentration was determined by Qubit 4.0 (Invitrogen, USA). RCA signal was received by an RT-qPCR instrument (Bio-Rad, USA).

AQA labeling reaction

ODNs (100 μM, 1 μL), AQA (10 mM in DMSO, 1 μL), PB buffer (1 M, pH 7.0, 5 μL), and 43 μL ddH_2O were mixed in a 1.5 mL microcentrifuge tube and incubated at 37°C for 2 h. The DNA products were obtained from the reaction mixture by ethanol precipitation.

Electrophoresis

The 20% denaturing, 8%, 10% native polyacrylamide gel (19:1 monomer to bis ratio), or 1.5% agarose gel electrophoresis were used to separate and analyze the reaction of nucleic acid. Electrophoresis was run in a temperature-controlled vertical and about 50 ng of DNA with different treatments were loaded on the gel. Electrophoresis was run at RT and the DNA was stained with Super GelRed dye. The final in-gel DNA was visualized by a Pharos FX Molecular imager (Bio-Rad, USA) in the UV mode. For the chain extension experiment, the gel was imaged directly in the 488 nm fluorescent mode without stain.

Chain extension experiment

The DNA 15 nt-5fU or 28 nt-5fU oligos were labeled with compound AQA as the method mentioned above to form 15 nt-5fU-AQA (EA) or 28 nt-5fU-AQA (EB). Then, in a 3.5 uL reaction system, 1 uM EA or EB chains, 1× phi29 DNA polymerase buffer (50 mM Tris-HCl, 10 mM MgCl2, 10 mM $(NH_4)_2SO_4$, 4 mM DTT, pH 7.5 at room temperature), 0.1 mM dNTPs, 15 nM P1 or P2 with 0.5 U phi29 DNA polymerase were incubated in a PCR EP tube at 30°C for 2 h. The enzyme was inactivated by adding a 3.3-fold quenching solution (95% formamide, 25mM EDTA at pH 8.0) and heating at 95°C for 10 minutes. The final products were characterized by 20% denaturing polyacrylamide gel electrophoresis at 350 V for 2.5 h at room

temperature.

Ligation reaction and product verification

In the 8-Strip PCR tube, 0.7 μM 80 bp template, 2.8 μM SL chain, 5.6 μL ddH_2O, and 1× Blunt/TA Ligase Master Mix were added and incubated at 25°C for 1 h in an RT-qPCR instrument. The reaction mixture was heated at 65°C for 10 min to inactivate the enzyme and stored at 4°C. The ligation product was verified by native 8% polyacrylamide gel electrophoresis at 240 V for 1.5 h.

The designed SL chain sequence contains a restriction enzyme Eam1104I cleavage recognition site CTCTTC (1/4), the ligation product could be confirmed based on this feature. By adding the secondary primer SPSL in the RCA process, double-stranded DNA containing restriction sites could be generated during the rolling circle amplification process. After the RCA reaction, the enzyme was inactivated at 65°C for 10 min, then slowly annealed to room temperature. By taking 2 μL of the above reaction mixture, 2 μL 10× FastDigest buffer and 1 μL Eam1104I, and 15 μL ddH_2O were added. The mixture was reacted overnight at 37°C and then at 80°C for 5 min to inactivate the enzyme, then slowly annealed to room temperature. The digested products were verified by 10% native polyacrylamide gel electrophoresis at 200 V for 2 h.

Chemical oxidation of ODNs

5hmU in ODNs were oxidized to 5fU based on a reported protocol.[3] Briefly, 1 μg 80 bp 5fU-DNA was made up to 24 μL containing 0.05 M NaOH on ice and then 1 μL $KRuO_4$ (15 mM in 0.05 M NaOH) was added to the mixture. The reaction was held on ice for about 1 h with occasional vortexing. The final reaction mixture was purified by a mini quick spin oligo column (BioRad, USA) to obtain the oxidized products.

Rolling circle amplification reaction

Briefly, a 20 μL solution system contained a 4 nM amplification template, 10 nM PSL, 1 U phi29 DNA polymerase, 1× phi29 DNA polymerase buffer, 0.25 mM dNTPs, and 1× SYBR Green II. Fluorescence signals were recorded every minute using an RT-qPCR instrument at 30°C for 30 min.

Quantitative analysis of genome-wide 5hmU and 5fU based on mass spectrometry

20 μg genomic fragments (23 μL) were heated at 95°C for 10 min and immediately placed on ice for 5 min. Then 6 μL 5× Nuclease S1 buffer and 10 U Nuclease S1 were added to the mixture and incubated at 37°C overnight. 3.5 μL 10× CIAP buffer, 1 U rSAP, and 1 μL VPDE (≥ 0.001 U) were in a final volume of 35 μL and digested at 37°C for 4 h. The final reaction mixture was mixed with 5 μL ddH2O and centrifuge at 12000 g for 10 min. 35 μL supernatant was used for LC-MS/MS to quantify 5hmC or 5fU in the whole genome.

To draw the standard working curve of 5hmU or 5fU, samples containing 2 μM 5hmU or 5fU monomers and 1 μM dNTPs were diluted to form a series of 50 μL solutions containing 2000, 1000, 500, 250, 125, 62.5 nM 5hmU or 5fU. All of the samples prepared for the LC-MS/MS were centrifuged at 12000 g for 10 min before injection.

Genomic DNA extraction

Genomic DNAs were from LO2 and MCF-7 cells and prepared by DNeasy Blood & Tissue Kit (QIAGEN) according to the manufacturer's instructions. The extracted DNAs were quantified via UV absorbance at

260 nm on a Nanodrop 2000c spectrophotometer and stored at −20°C.

Sample fragmentation

Genomic DNA was sheared into about 200 bp fragments on a Bioruptor® Sonication System Bioruptor® Plus with 30 s on and 30 s off for 35 cycles. The sample fragments were characterized by 1.5% agarose gel electrophoresis at 120 V for 30 min at room temperature and stored at −20°C.

Adaptor ligation for genomic DNA

Based on the manufacturer's instruction, 1 μg of the DNA fragments were repaired by reacting them with the 1× End Prep Enzyme Mix and 1× End Repair Reaction Buffer in a total of 10 μL reaction system at 20°C for 30 min and then at 65°C for 30 min. Subsequently, 4 μL reaction mixture above was reacted with the 5.6 μL 2× Blunt/TA Ligase Master Mix and 1.6 μL 10 μL SL at 25°C for 1 h and then 65°C for 10 min. The obtained adaptor-containing fragments were labeled with AQA following the method used earlier in the article to obtain chemically labeled fragments.

Labeling of epigenetic bases in genomic DNA fragments

The labeling with AQA and purification of the epigenetic bases in genomic DNA fragments was performed as the method above. The final concentrations of labeled DNA were evaluated via UV absorbance at 260 nm on a Nanodrop 2000c spectrophotometer.

Rolling circle amplification reaction with genomic DNA sample

The amplification reaction for biological samples was conducted at 37°C for 30 min in 20 μL solutions containing 40 ng different target fragments respectively (control DNA, 5fU DNA, 5hmU DNA, and 5fU-AQA DNA), 10 nM PSL, 1 U phi29 DNA polymerase, 1× phi29 DNA polymerase buffer, 250 nM dNTPs and 1× SYBR Green II. The fluorescence signals were recorded per minute using an RT-qPCR instrument.

Data processing

Data manipulation was performed in Origin. All the experimental data from each real-time fluorescence spectrum were analyzed using equations (1) and (2) to calculate the slopes of the amplification plot and the Pearson product-moment correlation coefficient.[4]

$$slope_{i,j} = \frac{\sum_i^j (x_i - \bar{x}_{i,j})(y_i - \bar{y}_{i,j})}{\sum_i^j (x_i - \bar{x}_{i,j})^2} \quad (1)$$

$$r_{i,j} = \frac{\sum_i^j (x_i - \bar{x}_{i,j})(y_i - \bar{y}_{i,j})}{\sqrt{\sum_i^j (x_i - \bar{x}_{i,j})^2 \sum_i^j (y_i - \bar{y}_{i,j})^2}} \quad (2)$$

$\bar{x}_{i,j}$ and $\bar{y}_{i,j}$ are sample averages, $slope_{ij}$ represent the slope of the linear regression equation from the i^{th} cycle to the j^{th} cycle, and r_{ij} represent the corresponding Pearson product-moment correlation coefficient. The square of the Pearson coefficient (r) is R^2.

Table S1. Synthetic ODN used in this study

.

Name	Sequence from 5’ to 3’
80-bp-template	TTCTTGGCTGTGGCTCTGCGTCCTTGTCCTGCCA CCCGCCCGGCGGGTGGAGGCACAACAGAGAGC AACACCGCCGAGGA
80-bp-template forward primer	P-TTCTTGGCTGTGGCTCTGCGTCCTTGTCCT
80-bp-template reverse primer	P-TCCTCGGCGGTGTTGCTCTCTGTTGTGCCT
15 nt-5fU	GACTCAA5fUAGCCGTA
15 nt-5fC	GACTCAA5fCAGCCGTA
15 nt-AP	GACTCAAAPAGCCGTA
15 nt-T	GACTCAATAGCCGTA
28 nt-5fU	TCGACTGTGTGTATC5fUGTGTGAGTCGAG
P1	FAM-TACGGTACGGCT
P2	FAM-CTCGACTCACAC
SL	P-CCTAGGGCCGCGCTCTCTTCACATACACCTCTCC TCCGCGCCCTAGGT
PSL	GGAGAGGTGTATGTGAAG
SPSL	CCTCCGCGGCCC

Table S2. Slope screening in the confidence interval with $R^2 > 0.9995$.

$Cycle_i$ \ $Cycle_j - Cycle_i$	5	6	7	8	9	10
5	0.9900	0.9931	0.9955	0.9967	0.9974	0.9977
6	0.9997	0.9972	0.9968	0.9977	0.9982	0.9979
7	0.9956	0.9959	0.9971	0.9975	0.9971	0.9961
8	0.9969	0.9968	0.9968	0.9959	0.9947	0.9927
9	0.9951	0.9950	0.9940	0.9928	0.9909	0.9926
10	0.9916	0.9916	0.9913	0.9900	0.9925	0.9928

$Cycle_i$ presents the initial screening cycle site; $Cycle_j$ presents the terminal cycle site.

Table S3. Precision and recoveries for the determination of 5fU by OFCRCA strategy.

Sample	Detected (pM)	Add (pM)	Found (pM)	Recovery	R.S.D
Sample 1	45.4	45.4	96.5	112.4%	1.2%
Sample 2	45.4	136.2	174.1	94.5%	1.0%
Sample 3	45.4	227	273.9	100.7%	0.9%

Table S4. Precision and recoveries for the determination of 5hmU by OFCRCA strategy.

Sample	Detected (pM)	Add (pM)	Found (pM)	Recovery	R.S.D
Sample 1	45.4	45.4	96.6	112.8%	2.9%
Sample 2	45.4	136.2	178.5	97.7%	7.4%
Sample 3	45.4	227	257.5	93.4%	3.3%

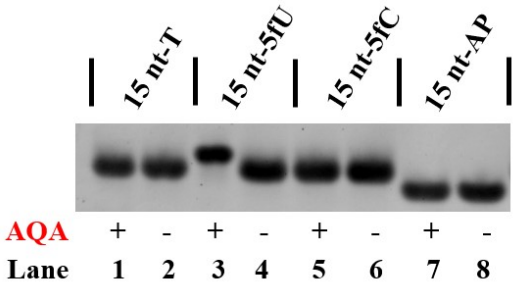

Figure S1. Polyacrylamide gel electrophoresis analysis of reaction between ODNs and AQA.

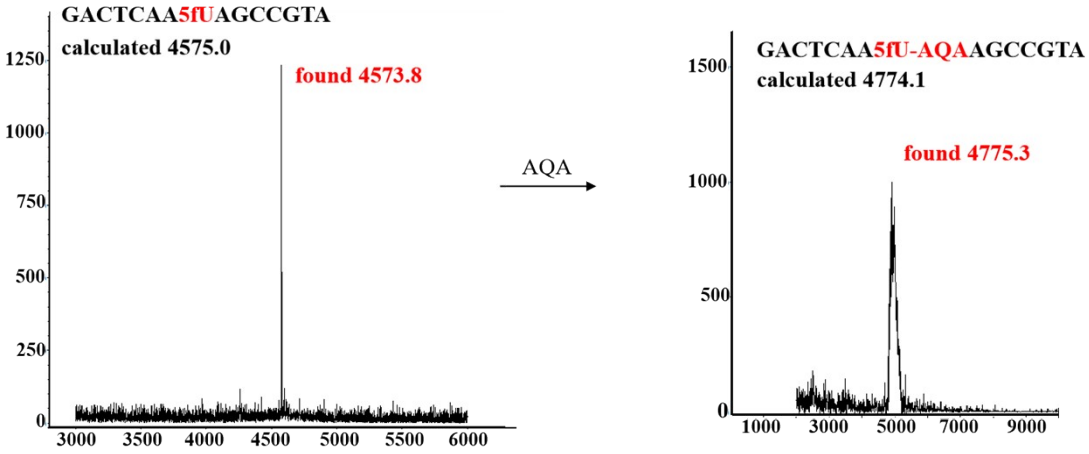

Figure S2. MALDI-TOF-MS spectrum of 5fU-DNA labeled with AQA.

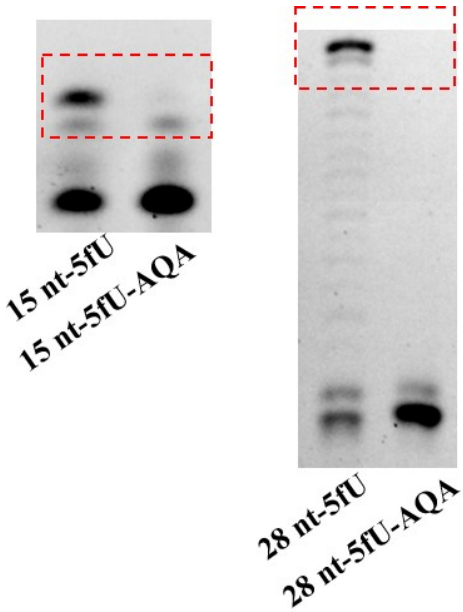

Figure S3. Polyacrylamide gel electrophoresis analysis of primer extension assay with phi29 DNA polymerase.

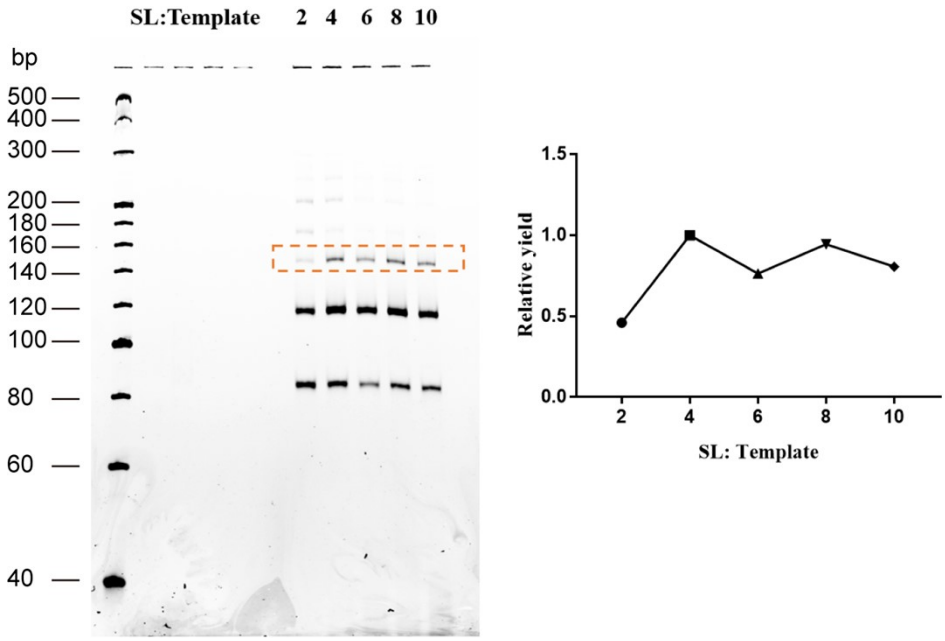

Figure S4. Optimization of the molar ratio of the DNA sample and adapters (SL) for ligation reaction. PAGE (polyacrylamide gel electrophoresis) analysis of the ligation reaction (left) and line chart analysis of the grey value from the PAGE.

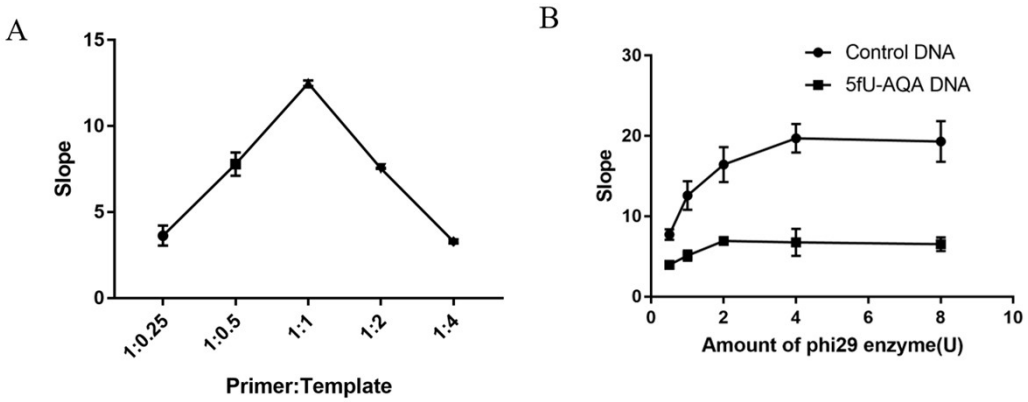

Figure S5. (**A**) Optimization of the molar ratio of the template and primer, the optimal molar ratio for the primer to the template was 1:1. (**B**) Optimization of the concentration of phi29 polymerase, the optimal concentration in the model reaction was 2U.

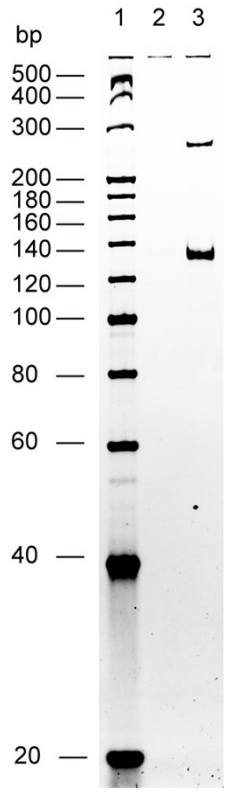

Figure S6. 10% PAGE analysis of digested products of RCA products. Lane 1 was the 20 bp DNA marker. Lane 2 was the digestion map of the RCA product using only the primary primer (PSL). The main product was single-stranded, which cannot be recognized by the enzyme cleavage site. Lane 3 was the digestion map of the RCA products using both primary primers (PSL) and secondary primers (SPSL). Bands of about 130 bp and 260 bp appeared, proving that the RCA proceeded successfully.

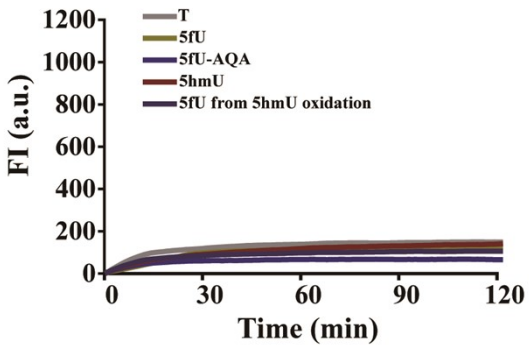

Figure S7. Time-dependent fluorescence spectra of the RCA reactions without adaptor ligation for T-DNA, 5fU-DNA, 5fU-AQA-DNA, 5hmU-DNA, and 5fU-DNA from 5hmU oxidation.

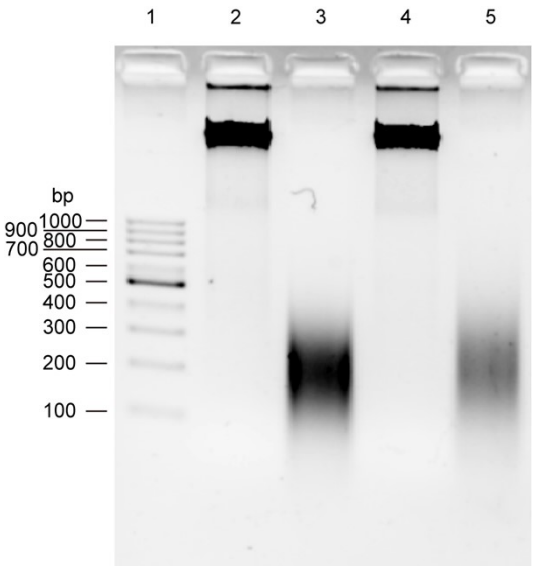

Figure S8. 1.5% agarose gel electrophoresis analysis of the fragmentation of whole-genome DNA. Lane 1 was the 100 bp DNA marker. Lanes 2 and 4 were the genomic DNA of LO2 and MCF-7 cells, respectively, and lanes 3 and 5 were approximately 200 bp fragments after ultrasonic shearing of the genomic DNA of LO2 and MCF-7 cells.

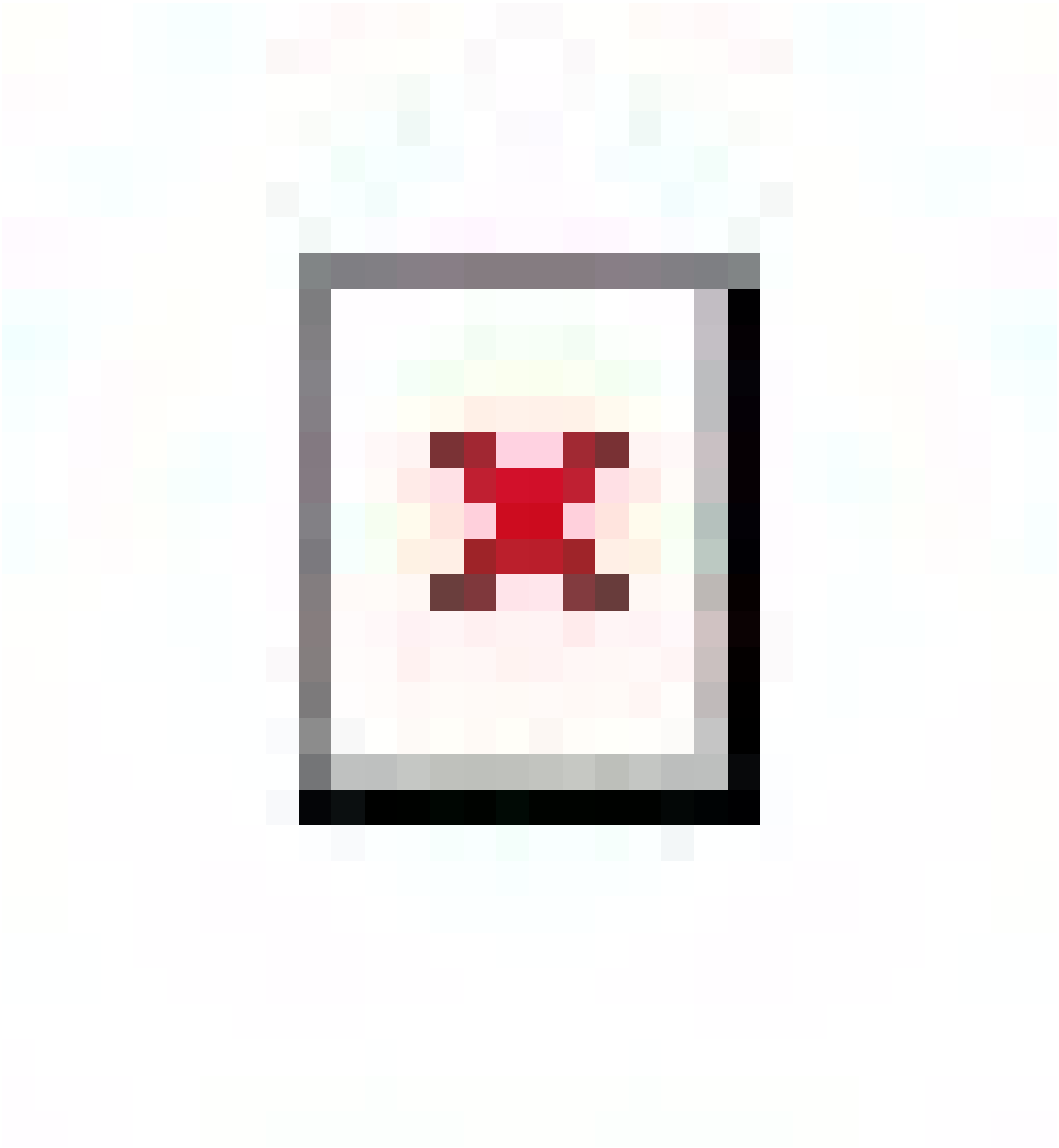

Figure S9. The standard curve was obtained by LC-MS/MS for quantitation of the total content of 5fU (**A**) and 5hmU (**B**) in MCF-7 cells. (**C**) Mass chromatograms of the quantifer transition for A, C, 5fU, and 5hmU standards. Mass chromatograms of the quantifer transition for LO2 (**D**) and MCF-7 (**E**) genomic DNAs. The content of 5fU was quantified to be 9.1 per10^6 dNs in LO2 genomic DNA and 10.3 per10^6 dNs in MCF genomic DNA.

References

[1] S. Wang, J. Wang, B. Fu, K. Chen, W. Xiong, L. Wei, G. Qing, T. Tian, X. Zhou, *J. Am. Chem. Soc.* **2018**, *140*, 15842-15849.

[2] J. Caton-Williams, M. Smith, N. Carrasco, Z. Huang, *Org. Lett.* **2011**, *13*, 4156-4159.

[3] R. E. Hardisty, F. Kawasaki, A. B. Sahakyan, S. Balasubramanian, *J. Am. Chem. Soc.* **2015**, *137*, 9270-9272.

[4] D. Chen, Y. Wang, M. Mo, J. Zhang, Y. Zhang, Y. Xu, S. Liu, J. Chen, Y. Ma, L. Zhang, Z. Dai, C. Cai, X. Zou, *Nucleic Acids Res.* **2019**, *47*, e119.