Electronic Supplementary Material (ESI) for Chemical Science. This journal is © The Royal Society of Chemistry 2022

Supporting Information

for

## Bisulfite-free and single-nucleotide resolution sequencing of DNA epigenetic

## modification of 5-hydroxymethylcytosine by engineered deaminase

Neng-Bin Xie,<sup>1,2,†</sup> Min Wang,<sup>1,†</sup> Tong-Tong Ji,<sup>1</sup> Xia Guo,<sup>1</sup> Jiang-Hui Ding,<sup>1</sup>

Bi-Feng Yuan,<sup>1,2,3,\*</sup> Yu-Qi Feng<sup>1,2</sup>

<sup>1</sup> Sauvage Center for Molecular Sciences, Department of Chemistry, Wuhan University,

Wuhan 430072, China.

<sup>2</sup> School of Public Health, Wuhan University, Wuhan 430071, China.

<sup>3</sup> Department of Radiation and Medical Oncology, Zhongnan Hospital of Wuhan University,

Wuhan 430071, China.

<sup>†</sup> These authors contributed equally to this work.

\* Corresponding author:

Bi-Feng Yuan. E-mail: bfyuan@whu.edu.cn

# **Table of Contents**

| Page S3  | Colony sequencing  |
|----------|--|
| Page S4  | Table S1. Sequences of oligonucleotides.   |
| Page S5  | Table S2. Sequences of double-stranded DNA.  |
| Page S6  | Table S3. Information of detected 5hmC sites in genomic DNA of human lung tissue.    |
| Page S7  | Figure S1. SDS-PAGE analysis of the purified wtA3A, eA3A-5 and eA3A-9.               |
| Page S8  | Figure S2. Principle of the AMD-seq method.  |
| Page S9  | Figure S3. Characterization of the deaminase selectivity of wtA3A toward C, 5mC and  |
|          | 5hmC in different sequence context by Sanger sequencing.                             |
| Page S10 | Figure S4. Characterization of the deaminase selectivity of eA3A-1 toward C, 5mC     |
|          | and 5hmC in different sequence context by Sanger sequencing.                         |
| Page S11 | Figure S5. Characterization of the deaminase selectivity of eA3A-2 toward C, 5mC     |
|          | and 5hmC in different sequence context by Sanger sequencing.                         |
| Page S12 | Figure S6. Characterization of the deaminase selectivity of eA3A-3 toward C, 5mC     |
|          | and 5hmC in different sequence context by Sanger sequencing.                         |
| Page S13 | Figure S7. Characterization of the deaminase selectivity of eA3A-4 toward C, 5mC     |
|          | and 5hmC in different sequence context by Sanger sequencing.                         |
| Page S14 | Figure S8. Characterization of the deaminase selectivity of eA3A-6 toward C, 5mC     |
|          | and 5hmC in different sequence context by Sanger sequencing.                         |
| Page S15 | Figure S9. Characterization of the deaminase selectivity of eA3A-7 toward C, 5mC     |
|          | and 5hmC in different sequence context by Sanger sequencing.                         |
| Page S16 | Figure S10. Characterization of the deaminase selectivity of eA3A-8 toward C, 5mC    |
|          | and 5hmC in different sequence context by Sanger sequencing.                         |
| Page S17 | Figure S11. Characterization of 5mC and 5hmC in CC site by EDM-seq analysis with     |
|          | Sanger sequencing.   |
| Page S18 | Figure S12. Extracted-ion chromatograms of dA, dG and dT from eA3A-5 or eA3A-9       |
|          | treated DNA by LC-MS/MS analysis.  |
| Page S19 | Figure S13. Evaluation of the deaminase activity of eA3A-5 to C, 5mC and 5hmC in     |
|          | GC and AC sites by colony sequencing.  |
| Page S20 | Figure S14. Evaluation of the deaminase activity of eA3A-9 to C, 5mC and 5hmC in     |
|          | TC and CC sites by colony sequencing.  |
| Page S21 | Figure S15. Evaluation of the deaminase activity of wtA3A to C, 5mC and 5hmC in      |
|          | GC and AC sites by colony sequencing.  |
| Page S22 | Figure S16. Evaluation of the deaminase activity of wtA3A to C, 5mC and 5hmC in      |
|          | TC and CC sites by colony sequencing.  |
| Page S23 | Figure S17. Quantitative evaluation of the level of 5hmC at AC and TC sites of DNA   |
|          | using EDM-seq with Sanger sequencing.  |
| Page S24 | Figure S18. Site-specific detection of 5hmC in genomic DNA of lung cancer tissue and |
|          | adjacent normal tissue by EDM-seq and AMD-seq.                                       |
| Page S25 | References   |

#### **Colony sequencing**

The colony sequencing was carried out according to the previous reports.<sup>1, 2</sup> Briefly, to quantitatively determine the ratio of C/(C+T) at given sites after the treatment of eA3A proteins, the deaminase-treated DNA was amplified and the PCR products were ligated into pClone007 Versatile Simple vector (Tsingke, Beijing, China). The resulting vector was then transformed into competent *E. coli* cells. Individual clone was randomly picked, lysed in TE buffer, amplified by PCR using M13 forward and reverse primers and the PCR products were then sequenced using an ABI3700 (Applied Biosystems, Inc.). Fifty positive clones per sample were picked and subjected to sequencing. The percentage of C/(C+T) at given sites could be calculated from the statistical analysis of fifty sequences, which reflect the percentage of 5hmC at given sites.

| Name    | Sequence (5' to 3')              |
|---------|----------------------------------|
| GC-C    | GTATGATGCGAATGAGATGTATTG         |
| AC-C    | GTATGATACGAATGAGATGTATTG         |
| TC-C    | GTATGATTCGAATGAGATGTATTG         |
| CC-C    | GTATGAT <u>C</u> GAATGAGATGTATTG |
| GC-5mC  | GTATGATG5mCGAATGAGATGTATTG       |
| AC-5mC  | GTATGATA5mCGAATGAGATGTATTG       |
| TC-5mC  | GTATGATT5mCGAATGAGATGTATTG       |
| CC-5mC  | GTATGATC5mCGAATGAGATGTATTG       |
| GC-5hmC | GTATGATG5hmCGAATGAGATGTATTG      |
| AC-5hmC | GTATGATA5hmCGAATGAGATGTATTG      |
| TC-5hmC | GTATGATT5hmCGAATGAGATGTATTG      |
| CC-5hmC | GTATGATC5hmCGAATGAGATGTATTG      |

 Table S1. Sequences of oligonucleotides.

Note: the "<u>C</u>" represented 5'-aza-2'-deoxycytidine.

DNA-C GAGTGACGCTGAGCTTGACGTCGCGCGATGAGAGGTGA TTATGAGTATGTATAGTGTTAGGAAGAGTGTAGTAATAG GATGAAGATGATTATATGATCGATGGTCCGTATGCGTAG AATACGTTGTTGTAGTGATTATAATGGAGTGAGAATGTA GATGAGTGGAGTAGGTAGTAGTAGATGTAGTGGTGAAGAG AGTAATTGTTAGTGGAATGTTGGAGAGGAT DNA-5mC GAGTGACGCTGAGCTTGACGTCGCGCGATGAGAGGTGA TTATGAGTATGTATAGTGTTAGGAAGAGTGTAGTAATAG GATGAAGATGATTATATGAT5mCGATGGT5mC5mCGTAT G5mCGTAGAATA5mCGTTGTTGTAGTGATTATAATGGAG TGAGAATGTAGATGAGTGGAGTAGGTAGTAAGATGTAG TGGTGAAGAGAGTAATTGTTAGTGGAATGTTGGAGAGG AT DNA-5hmC GAGTGACGCTGAGCTTGACGTCGCGCGATGAGAGGTGA TTATGAGTATGTATAGTGTTAGGAAGAGTGTAGTAATAG GATGAAGATGATTATATGAT5hmCGATGGT5hmC5hmCGT ATG5hmCGTAGAATA5hmCGTTGTTGTAGTGATTATAATG GAGTGAGAATGTAGATGAGTGGAGTAGGTAGTAAGATG TAGTGGTGAAGAGAGTAATTGTTAGTGGAATGTTGGAG AGGAT DNA-C5mC GAGTGACGCTGAGCTTGACGTCGCGCGCGTC5mCGAGT5mC 5mCGATGAGAGGTGATTATGAGTATGTATAGTGTTAGGA GGTTTGTATGTGTAGAATATGTTGTTGTAGTGATTATAA TGGAGTGAGAATGTAGATGAGTGGAGTAGGTAGTAAGA TGTAGTGGTGACAGAGAGTAATTGTTAGTGGAATGTTGG AGAGGAT GAGTGACGCTGAGCTTGACGTCGCGCGTC5hmCGAGT5h DNA-C5hmC mC5hmCGATGAGAGGTGATTATGAGTATGTATAGTGTTA GGAAGAGTGTAGTAATAGGATGAAGATGATTATATGAT TGATGGTTTGTATGTGTGTAGAATATGTTGTTGTAGTGATT ATAATGGAGTGAGAATGTAGATGAGTGGAGTAGGTAGT

Sequence (5' to 3')

Table S2. Sequences of double-stranded DNA.

Name

GTTGGAGAGGAT

AAGATGTAGTGGTGACAGAGAGTAATTGTTAGTGGAAT

| Site type | Primers for detection (from 5' to 3') |
|-----------|---------------------------------------|
| GC        | S: TTTAAATATTTTTTTAGGTGAATTATTTATG    |
|           | R: ACTATTACCTTTATCCTCCTCCCAT          |
| AC        | S: TGGGAAGATTGTTTGAGGTTAGG            |
|           | R: CATAAAACCACTCACTCTATCACCC          |
| TC        | S: GGTATTTATAGTAGTTTTGGAAGAATTTAGTG   |
|           | R: ACAAATATATAACACATCCATAAACACCT      |
| CC        | S: TAAATGTGGAGGAGGAATTAGATTTAT        |
|           | R: AAAATCACCTTTCCATCCCTACC            |
|           | Site type<br>GC<br>AC<br>TC<br>CC     |

**Table S3.** Information of detected 5hmC sites in genomic DNA of human lung tissue.

**Figure S1**. SDS-PAGE analysis of the purified wtA3A, eA3A-5 and eA3A-9. "M" represents protein marker. (A) SDS-PAGE analysis of the purified wtA3A. (B) SDS-PAGE analysis of the purified eA3A-5. (C) SDS-PAGE analysis of the purified eA3A-9. Line 1, without addition of IPTG; line 2, with addition of IPTG; line 3, purified proteins.



**Figure S2**. Principle of the AMD-seq method. (A) C, 5mC and 5hmC can be deaminated by wtA3A to form U, T and 5-hydroxymethyluracil (5hmU), respectively. All these deaminated products base pair with adenine. (B) In AMD-seq, DNA is firstly treated with  $\beta$ -GT that can transfer a glucosyl moiety to 5hmC to form 5gmC. Subsequent wtA3A treatment will lead to the deamination of C and 5mC to form U and T, both of which are read as T during sequencing. However, 5gmC is resistant to the deamination by wtA3A and then is still read as C during sequencing. Therefore, the remaining C manifests the original 5hmC site in DNA.



**Figure S3**. Characterization of the deaminase selectivity of wtA3A toward C, 5mC and 5hmC in different sequence context by Sanger sequencing. The 224-bp DNA-C, DNA-5mC and DNA-5hmC were used for the evaluation. All the C and 5mC were deaminated and then read as T. 5hmC in TC and CC sites were also deaminated and read as T. However, 5hmC in GC and AC sites were partially deaminated and therefore were partially read as C and partially read as T.

wtA3A DNA-C TCGATGGTCCGTATGCGTAGAATACG Sequenced TTGATGGT GT TGTAG ATG Т Т A TG DNA-5mCT GAT Sequenced T T G AT GGT GTAT G T G T A G A A T A T G Т Sm **DNA-5hmC** GATG GTAG Sequenced T GC G G G G С A G G G

**Figure S4**. Characterization of the deaminase selectivity of eA3A-1 toward C, 5mC and 5hmC in different sequence context by Sanger sequencing. The 224-bp DNA-C, DNA-5mC and DNA-5hmC were used for the evaluation. (A) The amino acid composition of wtA3A and eA3A proteins. (B) The sequencing results of DNA-C, DNA-5mC, and DNA-5hmC after eA3A-1 treatment. All the C were deaminated and read as T; but 5mC and 5hmC were partially deaminated and were partially read as C and partially read as T.

| (A)      |               |                                       |           |                      |  |
|----------|---------------|---------------------------------------|-----------|----------------------|--|
|          | α1            | Loop 1                                | β6        | Loop 7               |  |
| wtA3A    | 20<br>HIFTSNE | 30<br>NNGIGRHK                        | )<br>AAR] | 130<br>YDYDPL        |  |
| → eA3A-1 | HIFTSNE       | 'NNGIGR <mark>RQ</mark>               | AAR]      | YDYDPL               |  |
| eA3A-2   | HIFTSNE       | NNGIGR <mark>RQ</mark>                | AAR]      | YDYDTD               |  |
| eA3A-3   | HIFTSNE       | MNPIGR <mark>RQ</mark>                | AAR]      | YDYD <mark>TD</mark> |  |
| eA3A-4   | HIFTSNE       | NNDIGR <mark>RQ</mark>                | AAR]      | YDYDTD               |  |
| eA3A-5   | HIFTSNE       | MNNNIGR <mark>RQ</mark>               | AAR]      | YDYD <mark>TD</mark> |  |
| eA3A-6   | HIFTSNE       | TNNTIGR <mark>RQ</mark>               |           | YDYDTD               |  |
| eA3A-7   | HIFTSNE       | NNPI PRRQ                             | AARI      | YDYD <mark>TD</mark> |  |
| eA3A-8   | HIFTSNE       | TNNDIGR <mark>RQ</mark>               | AAR]      | YTYDTD               |  |
| eA3A-9   | HIFTSNE       | NN <mark>NN</mark> GR <mark>RQ</mark> | AAR]      | YDYDTD               |  |

 **Figure S5**. Characterization of the deaminase selectivity of eA3A-2 toward C, 5mC and 5hmC in different sequence context by Sanger sequencing. The 224-bp DNA-C, DNA-5mC and DNA-5hmC were used for the evaluation. (A) The amino acid composition of wtA3A and eA3A proteins. (B) The sequencing results of DNA-C, DNA-5mC, and DNA-5hmC after eA3A-2 treatment. In GC and AC sites, all the C and 5mC were deaminated and read as T; but all the 5hmC were resistant to deamination and were still read as C. In TC and CC sites, all the C and 5mC were deaminated and were partially read as C and partially read as T.

| (A)      | <b>A</b> |                        |            |      |                      |  |
|----------|----------|------------------------|------------|------|----------------------|--|
|          | α1       | Loop 1                 |            | β6   | Loop 7               |  |
| wtA3A    | HIFTSN   | FNNGIGF                | 30<br>2.HK | AARI | YDYDPL               |  |
| eA3A-1   | HIFTSN   | FNNGIGF                | RQ         | AARI | YDYDPL               |  |
| → eA3A-2 | HIFTSN   | FNNGIGF                | RRQ        | AARI | YDYDTD               |  |
| eA3A-3   | HIFTSN   | FNN <mark>P</mark> IGF | RRQ        | AARI | YDYD <mark>TD</mark> |  |
| eA3A-4   | HIFTSN   | FNNDIGF                | RRQ        | AARI | YDYDTD               |  |
| eA3A-5   | HIFTSN   | FNN <mark>N</mark> IGF | RQ         | AARI | YDYDTD               |  |
| eA3A-6   | HIFTSN   | FNNTIGF                | RQ         | AARI | YDYDTD               |  |
| eA3A-7   | HIFTSN   | FNN <mark>PIP</mark> F | RRQ        | AARI | YDYDTD               |  |
| eA3A-8   | HIFTSN   | FNNDIGF                | RRQ        | AARI | YTYDTD               |  |
| eA3A-9   | HIFTSN   | FNN <mark>NN</mark> GF | RRQ        | AARI | YDYDTD               |  |

| (B)                 |  | eA3     | A-2     |        |
|---------------------|--|---------|---------|--------|
| DNA-C<br>Sequence   | ; TCGATG<br>dttgatg                          | GTCCGTA | TGCGTAG | AATACG |
|                     | <u>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</u> | M       |         |        |
| DNA-5m0<br>Sequence |  | GTCCGTA |         |        |
| DNA-5hm<br>Sequence |  | GTCCGTA |         |        |
|                     | $\Lambda \Lambda \Lambda$                    | MAM     |         | ////// |

**Figure S6**. Characterization of the deaminase selectivity of eA3A-3 toward C, 5mC and 5hmC in different sequence context by Sanger sequencing. The 224-bp DNA-C, DNA-5mC and DNA-5hmC were used for the evaluation. (A) The amino acid composition of wtA3A and eA3A proteins. (B) The sequencing results of DNA-C, DNA-5mC, and DNA-5hmC after eA3A-3 treatment. All the C were deaminated and read as T; but 5mC were partially deaminated and were partially read as C and partially read as T; while 5hmC were resistant to deamination and were still read as C.

| (A)      |                               |   |  |  |  |  |
|----------|-------------------------------|---|--|--|--|--|
|          | α 1 Loop 1                    | β 6 Loop 7                              |  |  |  |  |
| wtA3A    | HIFTSNFNNGIGRHK               | AARIYDYDPL                              |  |  |  |  |
| eA3A-1   | HIFTSNFNNGIGR <mark>RQ</mark> | AARIYDYDPL                              |  |  |  |  |
| eA3A-2   | HIFTSNFNNGIGR <mark>RQ</mark> | AARIYDYD <mark>TD</mark>                |  |  |  |  |
| → eA3A-3 | HIFTSNFNNPIGR <mark>RQ</mark> | AARIYDYD <mark>TD</mark>                |  |  |  |  |
| eA3A-4   | HIFTSNFNNDIGRRQ               | AARIYDYDTD                              |  |  |  |  |
| eA3A-5   | HIFTSNFNNNIGRRQ               | AARIYDYD <mark>TD</mark>                |  |  |  |  |
| eA3A-6   | HIFTSNFNNTIGR <mark>RQ</mark> | AARIYDYD <mark>TD</mark>                |  |  |  |  |
| eA3A-7   | HIFTSNFNNPIPRRQ               | AARIYDYD <mark>TD</mark>                |  |  |  |  |
| eA3A-8   | HIFTSNFNNDIGR <mark>RQ</mark> | AARIY <mark>T</mark> YD <mark>TD</mark> |  |  |  |  |
| eA3A-9   | HIFTSNFNNNNGRRQ               | AARIYDYDTD                              |  |  |  |  |

**Figure S7**. Characterization of the deaminase selectivity of eA3A-4 toward C, 5mC and 5hmC in different sequence context by Sanger sequencing. The 224-bp DNA-C, DNA-5mC and DNA-5hmC were used for the evaluation. (A) The amino acid composition of wtA3A and eA3A proteins. (B) The sequencing results of DNA-C, DNA-5mC, and DNA-5hmC after eA3A-4 treatment. All the C were deaminated and read as T; but 5mC were partially deaminated and were partially read as C and partially read as T; while 5hmC were resistant to deamination and were still read as C.

| (A)      |        |                       |                    |      |                      |
|----------|--------|-----------------------|--------------------|------|----------------------|
|          | α1     | Loop 1                | 67                 | β6   | Loop 7               |
| wtA3A    | HIFTSN | IFNNGI                | 30<br>GRHK         | AARI | 130<br>YDYDPL        |
| eA3A-1   | HIFTSM | IFNNGI                | GR <mark>RQ</mark> | AARI | YDYDPL               |
| eA3A-2   | HIFTSN | IFNNGI                | GR <mark>RQ</mark> | AARI | YDYD <mark>TD</mark> |
| eA3A-3   | HIFTSM | IFNNPI                | GR <mark>RQ</mark> | AARI | YDYD <mark>TD</mark> |
| → eA3A-4 | HIFTSN | IFNN <mark>D</mark> I | GR <mark>RQ</mark> | AARI | YDYD <mark>TD</mark> |
| eA3A-5   | HIFTSM | IFNN <mark>N</mark> I | GR <mark>RQ</mark> | AARI | YDYD <mark>TD</mark> |
| eA3A-6   | HIFTSN | IFNN <mark>T</mark> I | GR <mark>RQ</mark> | AARI | YDYD <mark>TD</mark> |
| eA3A-7   | HIFTSM | IFNNPI                | PRRQ               | AARI | YDYD <mark>TD</mark> |
| eA3A-8   | HIFTSN | IFNN <mark>D</mark> I | GR <mark>RQ</mark> | AARI | YTYDTD               |
| eA3A-9   | HIFTSN | IFNN <mark>NN</mark>  | GR <mark>RQ</mark> | AARI | YDYDTD               |

| (B)      |           | eA3  | A-4                     |             |
|----------|-----------|--|-------------------------|-------------|
| DNA-     |           | GGTCCGTA   |                         | AATACG      |
| oequene  | *         | **   | *                       | *           |
|          | $\sim$    | hin  |                         | M           |
|          | Q         | çç   | Q                       | Q           |
|          |           |  | тесстас                 |             |
| Sequenc  | ed TIGATO | G G T T T G T A  |                         |             |
|          |           |  |                         | $\wedge$    |
|          | $\sim$    | head   |                         | adda        |
|          | D<br>M    |  | o<br>E                  | C<br>E      |
| DNA-5hr  |           | GGTCCGTA   | AT <mark>GC</mark> GTAG | BAATACG     |
| Sequence |           | G G T C C G T A  |                         | A A T A C G |
|          | . Λ       | • • • • •  |                         |             |
|          | MW        | $\langle \sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{$ | WWM                     | VMVN        |

**Figure S8**. Characterization of the deaminase selectivity of eA3A-6 toward C, 5mC and 5hmC in different sequence context by Sanger sequencing. The 224-bp DNA-C, DNA-5mC and DNA-5hmC were used for the evaluation. (A) The amino acid composition of wtA3A and eA3A proteins. (B) The sequencing results of DNA-C, DNA-5mC, and DNA-5hmC after eA3A-6 treatment. In GC and AC sites, all the C and 5mC were deaminated and read as T; but all the 5hmC were resistant to deamination and were still read as C. In TC and CC sites, all the C and 5mC were deaminated and were partially read as C and partially read as T.

| (A)      |             |                                      | Ţ    |                         |
|----------|-------------|--------------------------------------|------|-------------------------|
|          | <u>α1</u> L | oop 1                                | β6   | Loop 7                  |
| wtA3A    | HIFTSNEN    | 30<br>NGIGRHK                        | AARI | YDYDPL                  |
| eA3A-1   | HIFTSNFN    | NGIGR <mark>RQ</mark>                | AARI | YDYDPL                  |
| eA3A-2   | HIFTSNFN    | NGIGR <mark>RQ</mark>                | AARI | YDYD <mark>TD</mark>    |
| eA3A-3   | HIFTSNFN    | N <mark>P</mark> IGR <mark>RQ</mark> | AARI | YDYD <mark>TD</mark>    |
| eA3A-4   | HIFTSNFN    | NDIGR <mark>RQ</mark>                | AARI | YDYD <mark>TD</mark>    |
| eA3A-5   | HIFTSNFN    | N <mark>N</mark> IGR <mark>RQ</mark> | AARI | YDYD <mark>TD</mark>    |
| → eA3A-6 | HIFTSNFN    | N <mark>T</mark> IGR <mark>RQ</mark> | AARI | YDYD <mark>TD</mark>    |
| eA3A-7   | HIFTSNFN    | NPIPRRQ                              | AARI | YDYD <mark>TD</mark>    |
| eA3A-8   | HIFTSNFN    | NDIGR <mark>RQ</mark>                | AARI | Y <b>T</b> YD <b>TD</b> |
| eA3A-9   | HIFTSNEN    | NNNGRRQ                              | AARI | YDYDTD                  |

| (B)                 |   | eA3      | <b>A-6</b> |         |
|---------------------|---|----------|------------|---------|
| DNA-C<br>Sequence   |   | GGTCCGTA | TGCGTAG    | GAATACG |
|                     |   | hm       |            | MMA     |
| DNA-5m0<br>Sequence | CTCGATO   | GGTCCGTA |            | GAATACG |
|                     | $\mathcal{M}$   | hm       |            |         |
| DNA-5hm<br>Sequence |   | GGTCCGTA |            |         |
|                     | $\sim \sim $ | M        |            | MM      |

**Figure S9**. Characterization of the deaminase selectivity of eA3A-7 toward C, 5mC and 5hmC in different sequence context by Sanger sequencing. The 224-bp DNA-C, DNA-5mC and DNA-5hmC were used for the evaluation. (A) The amino acid composition of wtA3A and eA3A proteins. (B) The sequencing results of DNA-C, DNA-5mC, and DNA-5hmC after eA3A-7 treatment. In GC and AC sites, all the C were deaminated and read as T; but 5mC were partially deaminated and were partially read as C and partially read as T; while all the 5hmC were resistant to deamination and were still read as C. In TC and CC sites, all the C and 5mC were deaminated and then read as T; but all the 5hmC were resistant to deamination and were still read as C.

| (A)      |                  |                      |      |                         |
|----------|------------------|----------------------|------|-------------------------|
|          | <u>α1</u> Loo    | p 1                  | β6   | Loop 7                  |
| wtA3A    | 20<br>HIFTSNFNNG | 30<br>GIGRHK         | AARI | 30<br>YDYDPL            |
| eA3A-1   | HIFTSNFNNG       | GIGR <mark>RQ</mark> | AARI | YDYDPL                  |
| eA3A-2   | HIFTSNFNNG       | GIGR <mark>RQ</mark> | AARI | YDYD <mark>TD</mark>    |
| eA3A-3   | HIFTSNFNN        | PIGR <mark>RQ</mark> | AARI | YDYD <mark>TD</mark>    |
| eA3A-4   | HIFTSNFNNI       | IGR <mark>RQ</mark>  | AARI | YDYD <mark>TD</mark>    |
| eA3A-5   | HIFTSNFNN        | IGR <mark>RQ</mark>  | AARI | YDYD <mark>TD</mark>    |
| eA3A-6   | HIFTSNFNN        | IGR <mark>RQ</mark>  | AARI | YDYD <mark>TD</mark>    |
| → eA3A-7 | HIFTSNFNN        | PIPRRQ               | AARI | YDYD <mark>TD</mark>    |
| eA3A-8   | HIFTSNFNNI       | IGR <mark>RQ</mark>  | AARI | Y <b>T</b> YD <b>TD</b> |
| eA3A-9   | HIFTSNFNN        | INGRRQ               | AARI | YDYDTD                  |

| (B)                 |   | eA3A      | -7            |         |
|---------------------|---|-----------|---------------|---------|
| DNA-C<br>Sequence   |   | GGTCCGTAT |               | GAATACG |
|                     | ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ | h         |               |         |
| DNA-5m0<br>Sequence |   | GGTCCGTAT | GCGTAC        | GAATACG |
|                     |   |           |               |         |
| DNA-5hm<br>Sequence |   | GGTCCGTA  |               | GAATACG |
|                     | M                                       | MM        | $\mathcal{M}$ | M       |

**Figure S10**. Characterization of the deaminase selectivity of eA3A-8 toward C, 5mC and 5hmC in different sequence context by Sanger sequencing. The 224-bp DNA-C, DNA-5mC and DNA-5hmC were used for the evaluation. (A) The amino acid composition of wtA3A and eA3A proteins. (B) The sequencing results of DNA-C, DNA-5mC, and DNA-5hmC after eA3A-8 treatment. In GC and AC sites, all the C were deaminated and read as T; but 5mC were partially deaminated and were partially read as C and partially read as T; while all the 5hmC were resistant to deamination and were still read as C. In TC and CC sites, all the C and 5mC were deaminated and then read as T; but all the 5hmC were resistant to deamination and were still read as C.

| (A)      |         |                                       |                     |                          |  |
|----------|---------|---------------------------------------|---------------------|--------------------------|--|
|          | α1      | Loop 1                                | β6                  | Loop 7                   |  |
| wtA3A    | HIFTSNF | 30<br>NNGIGRHF                        | )<br>[ <b>AA</b> R] | 130<br>IYDYDPL           |  |
| eA3A-1   | HIFTSNF | NNGIGR <mark>RÇ</mark>                |                     | IYDYDPL                  |  |
| eA3A-2   | HIFTSNF | NNGIGR <mark>RÇ</mark>                |                     | IYDYD <mark>TD</mark>    |  |
| eA3A-3   | HIFTSNF | NN <mark>P</mark> IGR <mark>RÇ</mark> |                     | IYDYD <mark>TD</mark>    |  |
| eA3A-4   | HIFTSNF | NN <mark>D</mark> IGR <mark>RÇ</mark> |                     | IYDYD <mark>TD</mark>    |  |
| eA3A-5   | HIFTSNF | NN <mark>N</mark> IGR <mark>RÇ</mark> |                     | IYDYD <mark>TD</mark>    |  |
| eA3A-6   | HIFTSNF | NNTIGR <mark>RÇ</mark>                |                     | IYDYD <mark>TD</mark>    |  |
| eA3A-7   | HIFTSNF | NNPIPRRC                              |                     | IYDYD <mark>TD</mark>    |  |
| → eA3A-8 | HIFTSNF | NNDIGR <mark>RÇ</mark>                |                     | IY <b>T</b> YD <b>TD</b> |  |
| eA3A-9   | HIFTSNF | NN <mark>NN</mark> GR <mark>RÇ</mark> |                     | IYDYD <mark>TD</mark>    |  |

| (B)                      | eA3A-  | 8            |
|--------------------------|--|--------------|
| DNA-C T<br>Sequenced T   | CGATGGTCCGTAT(                               | GCGTAGAATACG |
| $\overline{\mathcal{N}}$ | <u>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</u> |              |
| DNA-5mCT(<br>Sequenced⊺  | <b>CGATGGTCCGTATC</b>                        |              |
| ×                        |  | LAWAMAA      |
| DNA-5hmCT<br>Sequenced⊤  |  |              |
| $\wedge$                 | MMM  | MMMM         |

**Figure S11**. Characterization of 5mC and 5hmC in CC site by EDM-seq analysis with Sanger sequencing. DNA-C5mC and DNA-C5hmC were used for the evaluation. After eA3A-9 treatment, 5mC sites in both C5mC and 5mC5mC were deaminated and read as T; while 5hmC sites in both C5hmC and 5hmC5hmC were not deaminated and still read as C.

5m0 5mC DNA-C5mC CGTCCGAGTCCGATGAGA Sequenced TGTTTGAGTTTGATGAGA

DNA-C5hmC CGTCCGAGTCCGATGAGA Sequenced TGTTCGAGTCCGATGAGA

**Figure S12**. Extracted-ion chromatograms of dA, dG and dT from eA3A-5 or eA3A-9 treated DNA by LC-MS/MS analysis. The mixture of GC-C and AC-C was treated with eA3A-5. Another mixture of TC-C and CC-C was treated with eA3A-9. (A) Extracted-ion chromatograms of dA, dG and dT from GC-C and AC-C mixture without or with eA3A-5 treatment. (B) Extracted-ion chromatograms of dA, dG and dT from TC-C and CC-C mixture without or with eA3A-9 treatment.



**Figure S13**. Evaluation of the deaminase activity of eA3A-5 to C, 5mC and 5hmC in GC and AC sites by colony sequencing. The 224-bp DNA-C, DNA-5mC and DNA-5hmC were first denatured to ssDNA and then separately treated with eA3A-5 followed by colony sequencing. Fifty clones for each sample were randomly picked up and sequenced. (A) C in GC and AC sites from DNA-C were all deaminated and read as T. (B) 5mC in GC and AC sites from DNA-5mC were all deaminated and read as T. (C) 5hmC in GC and AC sites from DNA-5hmC were resistant to deamination and were still read as C. Blue stick represents T; Orange stick represents C.

| (A)        |             |
|------------|-------------|
| GC         | C AC        |
| Clone 1 -  |             |
| Clone 2 💳  | <b>—</b>    |
| Clone 3 -  | <b>—</b> —  |
| Clone 4 –  | <b>—</b> —  |
| Clone 5 —  | <b>—</b> —  |
| Clone 6 –  | <b>—</b> —  |
| Clone 7 🗕  | <del></del> |
| Clone 8 –  | <b>—</b> —  |
| Clone 9 -  | <b>—</b> —  |
| Clone 10   | <b>— —</b>  |
| Clone 11 - | <b>—</b> —  |
| Clone 12   | <b>—</b>    |
| Clone 13   | <b>—</b>    |
| Clone 14   | <b>— —</b>  |
| Clone 15   | <u> </u>    |
| Clone 16—  | <u> </u>    |
| Clone 17   | <u> </u>    |
| Clone 18   | <u> </u>    |
| Clone 19   |             |
| Clone 20   |             |
| Clone 21   |             |
| Clone 22   |             |
| Clone 22   |             |
| Clone 24   |             |
| Clone 25   |             |
| Clone 26   |             |
| Clone 27   |             |
| Clone 27   |             |
| Clone 28   |             |
| Clone 29   |             |
| Clone 30   |             |
|            |             |
| Clone 32   |             |
| Clone 33   |             |
| Clone 34   |             |
| Clone 35   |             |
| Clone 36   |             |
| Clone 37 — |             |
| Clone 38—  |             |
| Clone 39—  |             |
| Clone 40 — | <b>—</b> —  |
| Clone 41 — | <b>— —</b>  |
| Clone 42   | <b>—</b> —  |
| Clone 43 — | <b>—</b>    |
| Clone 44 — | <b>—</b> —  |
| Clone 45   | <b>—</b> —  |
| Clone 46   | <b>—</b> —  |
| Clone 47 — | <b>—</b> —  |
| Clone 48   | <b>→</b>    |
| Clone 49   | <b>—</b>    |
| Clone 50   | <b>—</b> —  |
| _          |             |
|            | 120         |

| (B)      | -5mC | –<br>5mC |          |
|----------|------|----------|----------|
| Clone 1  | GC   | AC       |          |
| Clone 2  | 4    |          |          |
| Clone 3  | -+   |          | <u> </u> |
| Clone 4  | -+   |          |          |
| Clone 5  | +    |          |          |
| Clone 6  | +    |          |          |
| Clone 7  | +    |          |          |
| Clone 0  | -    | _        |          |
| Clone 10 |      |          |          |
| Clone 11 |      |          |          |
| Clone 12 | 4    |          |          |
| Clone 13 | 4    |          |          |
| Clone 14 | 4    |          |          |
| Clone 15 | -    |          |          |
| Clone 16 | -+   |          |          |
| Clone 17 | -+   | _        |          |
| Clone 18 | +    | _        | ———      |
| Clone 19 | +    |          |          |
| Clone 20 | +    |          |          |
| Clone 21 | -    |          |          |
| Clone 22 |      |          |          |
| Clone 24 |      |          |          |
| Clone 25 | _    |          |          |
| Clone 26 | 4    |          |          |
| Clone 27 | -    |          |          |
| Clone 28 | -+   | _        | <u> </u> |
| Clone 29 | -+   | _        |          |
| Clone 30 | +    |          | <u> </u> |
| Clone 31 | +    |          |          |
| Clone 32 | +    |          |          |
| Clone 34 | -    |          |          |
| Clone 35 |      |          |          |
| Clone 36 |      |          |          |
| Clone 37 |      |          |          |
| Clone 38 | 4    |          |          |
| Clone 39 | 4    |          |          |
| Clone 40 | -    |          |          |
| Clone 41 | -+   |          | <u> </u> |
| Clone 42 | +    |          | ———      |
| Clone 43 | +    |          |          |
| Clone 44 | +    |          |          |
| Clone 45 | +    |          |          |
| Clone 46 |      |          |          |
| Clone 42 |      |          |          |
| Clone 40 |      |          |          |
| Clone 50 |      |          |          |
| Sione Ju |      |          |          |
|          |      | 12       | 0        |

|          | ç  | 2 2 |          |
|----------|----|-----|----------|
| (C)      | 44 | 2hu |          |
|          | G  |     |          |
| Clone 1  | -  |     |          |
| Clone 2  |    |     |          |
| Clone 3  |    |     |          |
| Clone 5  |    |     |          |
| Clone 6  |    |     |          |
| Clone 7  |    |     |          |
| Clone 8  | _  |     |          |
| Clone 9  | _  |     |          |
| Clone 10 | -  |     |          |
| Clone 11 | -  |     | <u> </u> |
| Clone 12 | -  |     | <u> </u> |
| Clone 13 | -  |     | <u> </u> |
| Clone 14 | -  |     | <u> </u> |
| Clone 15 | -  |     | <u> </u> |
| Clone 16 | -  |     |          |
| Clone 17 |    |     |          |
| Clone 18 |    |     |          |
| Clone 19 |    |     |          |
| Clone 20 |    |     |          |
| Clone 21 |    |     |          |
| Clone 22 |    |     |          |
| Clone 24 | _  |     |          |
| Clone 25 | _  |     |          |
| Clone 26 | _  |     |          |
| Clone 27 | -  |     |          |
| Clone 28 | -  |     | <u> </u> |
| Clone 29 | -  |     | <u> </u> |
| Clone 30 | -  |     | <u> </u> |
| Clone 31 | -  |     | <b>—</b> |
| Clone 32 | -  |     | <u> </u> |
| Clone 33 | -  |     | <u> </u> |
| Clone 34 | -  |     |          |
| Clone 35 |    |     |          |
| Clone 36 |    |     |          |
| Clone 38 |    |     |          |
| Clone 39 |    |     |          |
| Clone 40 |    |     |          |
| Clone 41 |    |     |          |
| Clone 42 | _  |     |          |
| Clone 43 | _  |     |          |
| Clone 44 | _  |     |          |
| Clone 45 | -  |     | <u> </u> |
| Clone 46 | -  |     | <u> </u> |
| Clone 47 | -  |     | <u> </u> |
| Clone 48 | -  |     | <u> </u> |
| Clone 49 | -  |     | <u> </u> |
| Clone 50 |    |     | <b>—</b> |
|          |    |     |          |

120

Figure S14. Evaluation of the deaminase activity of eA3A-9 to C, 5mC and 5hmC in TC and CC sites by colony sequencing. The 224-bp DNA-C, DNA-5mC and DNA-5hmC were first denatured to ssDNA and then separately treated with eA3A-9 followed by colony sequencing. Fifty clones for each sample were randomly picked up and sequenced. (A) C in TC and CC sites from DNA-C were all deaminated and read as T. (B) 5mC in TC and CC sites from DNA-5mC were all deaminated and read as T. (C) 5hmC in TC and CC sites from DNA-5hmC were resistant to deamination and were still read as C. Blue stick represents T; Orange stick represents C.

(A)

|            | тс | cc  |
|------------|----|-----|
| Clone 1 -  |    |     |
| Clone 2    |    |     |
| Clone 3    |    |     |
| Clone 4 -  |    |     |
| Clone 5 -  |    |     |
| Clone 6 -  |    |     |
| Clone 7 -  |    |     |
| Clone 8 -  |    |     |
| Clone 9 -  |    |     |
| Clone 10 - |    |     |
| Clone 11 - |    |     |
| Clone 12 - |    |     |
| Clone 13 - |    |     |
| Clone 14   |    |     |
| Clone 15   |    |     |
| Clone 16   |    |     |
| Clone 17   |    |     |
| Clone 18   |    |     |
| Clone 19   |    |     |
| Clone 20 - |    |     |
| Clone 21   |    |     |
| Clone 22 - |    |     |
| Clone 23   |    |     |
| Clone 24   |    |     |
| Clone 25   |    |     |
| Clone 20   |    |     |
| Clone 28   |    |     |
| Clone 29   | _  |     |
| Clone 29 - |    |     |
| Clone 31   | _  |     |
| Clone 32   | _  |     |
| Clone 33   | _  |     |
| Clone 34   | _  |     |
| Clone 35   |    |     |
| Clone 36   |    |     |
| Clone 37   | _  |     |
| Clone 38   |    |     |
| Clone 39   |    |     |
| Clone 40   |    |     |
| Clone 41   | _  |     |
| Clone 42 - | _  |     |
| Clone 43   | _  |     |
| Clone 44   | _  |     |
| Clone 45   |    |     |
| Clone 46   |    |     |
| Clone 47   |    |     |
| Clone 48   |    |     |
| Clone 49   |    |     |
| Clone 50 - | _  |     |
|            | _  |     |
|            |    | 100 |

| (B)      |   | -5mC |   |
|----------|---|------|---|
| Clone 1  |   |      |   |
| Clone 2  |   |      |   |
| Clone 2  | _ | _    |   |
| Clone 4  |   |      |   |
| Clone F  |   |      |   |
| Clone 5  |   |      |   |
| Clone 7  |   |      |   |
| Clone /  |   |      |   |
| Clone 8  |   |      |   |
| Clone 9  | _ |      |   |
| Clone 10 | _ |      |   |
| Clone 11 | _ |      |   |
| Clone 12 | _ |      |   |
| Clone 13 | _ |      |   |
| Clone 14 | _ |      |   |
| Clone 15 |   |      |   |
| Clone 16 | _ |      |   |
| Clone 17 | _ |      |   |
| Clone 18 | _ | _    |   |
| Clone 10 |   |      |   |
| Clone 20 |   |      |   |
| Clone 20 |   |      |   |
| Clone 21 |   |      |   |
| Clone 22 |   |      |   |
| Clone 23 | _ |      |   |
| Clone 24 | _ |      |   |
| Clone 25 | _ |      |   |
| Clone 26 | _ |      | _ |
| Clone 27 | _ |      |   |
| Clone 28 | _ |      |   |
| Clone 29 | _ |      |   |
| Clone 30 | _ |      |   |
| Clone 31 |   |      |   |
| Clone 32 |   |      |   |
| Clone 33 |   |      |   |
| Clone 34 |   |      |   |
| Clone 34 |   |      |   |
| Clone 35 |   |      |   |
| Clone 36 |   |      |   |
| Clone 37 | _ |      |   |
| Clone 38 | - |      |   |
| Clone 39 | _ |      | _ |
| Clone 40 | _ |      |   |
| Clone 41 |   |      |   |
| Clone 42 |   |      |   |
| Clone 43 | _ |      |   |
| Clone 44 |   |      |   |
| Clone 45 |   |      |   |
| Clone 46 |   |      |   |
| Clone 47 |   |      |   |
| Clone 47 |   |      |   |
|          |   |      |   |
| Cione 49 |   |      |   |
| Clone 50 | _ |      |   |
|          |   | 100  |   |

| (C)        | ShmC | Shmc      |
|------------|------|-----------|
|            | T    | LI.       |
| Clone 1    |      | - ii      |
| Clone 2    | -    |           |
| Clone 3    | -    |           |
| Clone 4    | -    |           |
| Clone 5    |      |           |
| Clone 6    |      |           |
| Clone 7    |      | <u>ii</u> |
| Clone 8    |      |           |
| Clone 9    | - 1  |           |
| Clone 10   | - 1  |           |
| Clone 11   | - 1  |           |
| Clone 12   | - 1  |           |
| Clone 12   |      |           |
| Clone 14   |      |           |
| Clone 14   |      |           |
| Clone 15   |      |           |
| Clone 16   |      |           |
| Clone 17   | -    |           |
| Clone 18   | -    |           |
| Clone 19   | -    |           |
| Clone 20 - | -    |           |
| Clone 21   | -    |           |
| Clone 22   | -    |           |
| Clone 23   | -    |           |
| Clone 24   | -    |           |
| Clone 25   | -    |           |
| Clone 26   | -    |           |
| Clone 27   | -    |           |
| Clone 28   | -    |           |
| Clone 29   | -    |           |
| Clone 30   |      |           |
| Clone 31   |      |           |
| Clone 32   |      |           |
| Clone 33   |      |           |
| Clone 34   | - 1  |           |
| Clone 35   | - 1  |           |
| Clone 36   | - 1  |           |
| Clone 27   |      |           |
| Clone 37   |      |           |
| Clone 30   |      |           |
| Clone 39   |      |           |
| Clone 40   |      |           |
| Clone 41   | -    |           |
| Clone 42   | -    |           |
| Clone 43   | -    |           |
| Clone 44   | -    |           |
| Clone 45   | -    |           |
| Clone 46   | -    |           |
| Clone 47   | -    |           |
| Clone 48   | -    |           |
| Clone 49   | -    |           |
| Clone 50   | -    |           |
|            |      |           |

100

**Figure S15.** Evaluation of the deaminase activity of wtA3A to C, 5mC and 5hmC in GC and AC sites by colony sequencing. The 224-bp DNA-C, DNA-5mC and DNA-5hmC were first denatured to ssDNA and then separately treated with wtA3A followed by colony sequencing. Fifty clones for each sample were randomly picked up and sequenced. (A) C in GC and AC sites from DNA-C were all deaminated and read as T. (B) 5mC in GC and AC sites from DNA-5mC were all deaminated and read as T. (C) 5hmC in GC and AC sites from DNA-5hmC were partially deaminated, and therefore were partially read as T and partially read as C. Blue stick represents T; Orange stick represents C.

| (A)      | (B) <sup>C</sup><br><sup>E</sup><br>I I | (C) 1 1 (C) |
|----------|---|-------------|
|          |   |             |
|          |   |             |
| Clone 2  | Clone 2                                 | Clone 2     |
| Clone 4  | Clone 4                                 |             |
| Clone 5  | Clone 5                                 |             |
| Clone 6  | Clone 6                                 | Clone 6     |
| Clone 7  | Clone 7                                 |             |
| Clone 8  | Clone 8                                 |             |
|          | Clone 9                                 |             |
| Clone 10 | Clone 10                                |             |
| Clone 11 |   | Clone 11    |
| Clone 12 | Clone 12                                | Clone 12    |
| Clone 13 | Clone 13                                | Clone 13    |
| Clone 14 | Clone 14                                | Clone 14    |
| Clone 15 | Clone 15                                | Clone 15    |
| Clone 16 | Clone 16                                | Clone 16    |
| Clone 17 | Clone 17                                | Clone 17    |
| Clone 18 | Clone 18                                | Clone 18    |
| Clone 19 | Clone 19                                | Clone 19    |
| Clone 20 | Clone 20                                | Clone 20    |
| Clone 21 | Clone 21                                | Clone 21    |
| Clone 22 | Clone 22                                | Clone 22    |
| Clone 23 | Clone 23                                | Clone 23    |
| Clone 24 | Clone 24                                | Clone 24    |
| Clone 25 | Clone 25                                | Clone 25    |
| Clone 26 | Clone 26                                | Clone 26    |
| Clone 27 | Clone 27                                | Clone 27    |
| Clone 28 |   |             |
| Clone 29 | Clone 29                                | Clone 29    |
| Clone 30 |   | Clone 30    |
| Clone 31 |   |             |
|          | Clone 32                                | Clone 32    |
| Clone 33 | Clone 34                                | Clone 34    |
| Clone 34 | Clone 35                                | Clone 35    |
| Clone 35 | Clone 36                                |             |
| Clone 27 | Clone 37                                | Clone 37    |
| Clone 37 |   |             |
| Clone 39 | Clone 39                                |             |
| Clone 40 | Clone 40                                |             |
| Clone 41 | Clone 41                                |             |
| Clone 42 | Clone 42                                | Clone 42    |
| Clone 43 |   | Clone 43    |
| Clone 44 | Clone 44                                | Clone 44    |
| Clone 45 | Clone 45                                | Clone 45    |
| Clone 46 | Clone 46                                | Clone 46    |
| Clone 47 | Clone 47                                | Clone 47    |
| Clone 48 | Clone 48                                | Clone 48    |
| Clone 49 | Clone 49                                | Clone 49    |
| Clone 50 | Clone 50                                | Clone 50    |
|          |   |             |
| 120      | 120                                     | 120         |

S21

**Figure S16.** Evaluation of the deaminase activity of wtA3A to C, 5mC and 5hmC in TC and CC sites by colony sequencing. The 224-bp DNA-C, DNA-5mC and DNA-5hmC were first denatured to ssDNA and then separately treated with wtA3A followed by colony sequencing. Fifty clones for each sample were randomly picked up and sequenced. (A) C in TC and CC sites from DNA-C were all deaminated and read as T. (B) 5mC in TC and CC sites from DNA-5mC were all deaminated and read as T. (C) 5hmC in TC and CC sites from DNA-5hmC were all deaminated and read as T. Blue stick represents T; Orange stick represents C.

| (A)            | (B) ម្ដី ម្ដី ម្ដី ម្ដី ម្ដី ម្ដី ម្ដី ម្ដី | (C)      |
|----------------|---|----------|
| TC CC          | TC CC                                       | τς τς ςς |
| Clone 1        |   |          |
| Clone 2        |   | Clone 2  |
| Clone 3        |   | Clone 3  |
| Clone 4        |   |          |
| Clone 5        |   | Clone 5  |
|                | Clone 7                                     |          |
|                |   |          |
| Clone 8        |   |          |
| Clone 9        | Clone 10                                    |          |
|                |   | Clone 11 |
|                |   | Clone 12 |
|                |   | Clone 13 |
|                |   | Clone 14 |
|                | Clone 15                                    | Clone 15 |
| Clone 15       | Clone 16                                    | Clone 16 |
|                | Clone 17                                    | Clone 17 |
|                | Clone 18                                    | Clone 18 |
| Clone 18       | Clone 19                                    | Clone 19 |
|                | Clone 20                                    | Clone 20 |
|                |   | Clone 21 |
|                |   | Clone 22 |
|                |   | Clone 23 |
|                | Clone 24                                    | Clone 24 |
| Clone 25       | Clone 25                                    | Clone 25 |
| Clone 26       | Clone 26                                    | Clone 26 |
|                | Clone 27                                    | Clone 27 |
|                | Clone 28                                    | Clone 28 |
|                | Clone 29                                    | Clone 29 |
|                | Clone 30 — H                                | Clone 30 |
|                | Clone 31                                    | Clone 31 |
|                | Clone 32                                    | Clone 32 |
|                | Clone 33                                    | Clone 33 |
| Clone 34       | Clone 34 — H                                | Clone 34 |
| Clone 35       | Clone 35                                    | Clone 35 |
| Clone 36       | Clone 36                                    | Clone 36 |
| Clone 37       | Clone 37                                    | Clone 37 |
| Clone 38 — III | Clone 38                                    | Clone 38 |
| Clone 39       | Clone 39                                    | Clone 39 |
| Clone 40       |   | Clone 40 |
| Clone 41 — H   |   | Clone 41 |
| Clone 42       |   | Clone 42 |
| Clone 43 — H   |   | Clone 43 |
| Clone 44 — H   |   | Clone 44 |
| Clone 45       |   | Clone 45 |
| Clone 46       |   | Clone 46 |
|                |   |          |
| Clone 48       |   |          |
| Clone 49       |   | Clone 49 |
| Clone 50       |   | Cione 50 |
|                | 100   |          |
| 100            | 100   | 100      |

**Figure S17**. Quantitative evaluation of the level of 5hmC at AC and TC sites of DNA using EDM-seq with Sanger sequencing. (A) DNA-C and DNA-5hmC were mixed at different ratios with DNA-5hmC ranging from 0% to 100%. The mixtures were treated with eA3A-5 followed by Sanger sequencing. (B) Linear regression of the measured percentage versus theoretical percentage of 5hmC at AC site by EDM-seq with Sanger sequencing. (C) DNA-C and DNA-5hmC were mixed at different ratios with DNA-5hmC ranging from 0% to 100%. The mixtures were treated with eA3A-9 followed by Sanger sequencing. (D) Linear regression of the measured percentage versus theoretical percentage of 5hmC at AC site by EDM-seq with Sanger sequencing. (D) Linear regression of the measured percentage versus theoretical percentage of 5hmC at TC site by EDM-seq with Sanger sequencing.



Figure S18. Site-specific detection of 5hmC in genomic DNA of lung cancer tissue and adjacent normal tissue by EDM-seq and AMD-seq. Shown in (A) and (B) are two replicates related to Figure 7.

| (A)           |   | Cancer tissue  | (B) Normal tissue  | Cancer tissue  |
|---------------|---|--|--|--|
| EDM-se        |   | EDM-seq A G A A A G A G G GT G AT A T T A A  | EDM-seq A G A A A G A G G G G G AT A T T A A   | EDM-seq A G A A A G A G G G T G A T A T T A A  |
|               | Marin   | <u>xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx</u>   | <u>Manna</u>   | aman   |
| AMD-S         | eq A G A A A G A G G <mark>C</mark> G A T A T T A A   | AMD-Seq A G A A A G A G G G T G A T A T T A A  | AMD-Seq A G A A A G A G G T G A T A T T A A  | AMD-Seq A G A A A G A G G T G A T A T T A A  |
|               | xxml Maxm   | <u>namalma</u> taam  | MMM MMM  | <u> </u>   |
|               |   |  |  |  |
|               |   |  |  |  |
| EDM-s         | eq G T G G T G G T A C G T A G T T G T.   | EDM-seqGTGGTGGTGGTATGTAGTTGT   | EDM-seqGTGGTGGTACGTAGTTGT  | EDM-seq GT G GT G GT A G TT GT   |
|               | <u>Samana (antana</u>   | <u>mmm</u>   | <u>Mamaa Mamaa</u>   | <u>Anna Manna</u>  |
| AMD-S         | <b>Seq</b> G T G G T G G T A G T T G T <b>A</b> G T A G T T G T   | AMD-Seq 6 T 6 6 T 6 6 T A 7 7 6 T A 6 T T 6 T  | АМ <b>D-Seq</b> 6 Т 6 6 Т 6 6 Т А 6 Т 7 6 Т 6 Т 6 Т 6 Т 6 Т 6 Т 6 Т 6 Т 6  | <b>AMD-Seq</b> G T G G T G G T A G T T G T A G T T G T   |
|               | Ama Anna  | <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>   | annahanna  | Amon Anna  |
|               |   |  |  |  |
|               |   |  |  |  |
|               | Normal tissue   | Cancer tissue  | Normal tissue  | Cancer tissue  |
| EDM-          | Normal tissue<br>GTGCCTCATCGTTTTTGT<br>seq 6 1 6 1 1 1 1 A 1 1 6 1 1 1 1 6 1  | Cancer tissue<br>GTGCCTCATCGTTTTTGT<br>EDM-seq 6 1 6 1 1 1 1 A 1 6 1 1 1 1 1 6 1   | Normal tissue<br>GTGCCTCATCGTTTTTGT<br>EDM-seq GT GT TTTAT   | Cancer tissue<br>GTGCCTCATCGTTTTTGT<br>EDM-seqG1 GIIIIAI <mark>I</mark> GIIIIIGI   |
| EDM-          | Normal tissue<br>GTGCCTCATCGTTTTTGT<br>seq 6 1 6 1 1 1 1 1 6 1 1 1 1 6 1<br>MMMMMMMMM   | Cancer tissue<br>GTGCCTCATCGTTTTTGT<br>EDM-seq 6 1 6 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1   | Normal lissue<br>GTGCCTCATCGTTTTTGT<br>EDM-seq GT GTTTTTTGT<br>AMMMMMM   | Cancer tissue<br>GTGCCTCATCGTTTTGT<br>EDM-segGT GTTTTAT_GTTTTTGT   |
| EDM-          |   |  |  |  |
| EDM-          |   | Cancer tissue<br>GTGCCTCATCGTTTTTGT<br>EDM-seq 6 1 6 1 1 1 1 4 1 6 1 1 1 1 6 1<br>AMD-Seq 6 1 6 1 1 1 1 4 1 6 1 1 1 1 6 1<br>AMD-Seq 6 1 6 1 1 1 1 4 1 6 1 4 1 1 1 6 1   |  |  |
| EDM-          | Normal tissue<br>GTGCCTCATCGTTTTTGT<br>seq 6 1 6 1 1 1 1 1 6 1 1 1 1 6 1<br>Seq 6 1 6 1 1 1 1 4 1 6 1 1 1 1 6 1<br>Seq 6 1 6 1 1 1 1 4 1 6 1 1 1 1 6 1<br>Address Address Add | Cancer tissue GTGCCTCATCGTTTTTGT EDM-seq GTGTTTTTATIGTTTTGT AMD-Seq GTGTTTTATIGTTTGT   |  |  |
| EDM-          | Normal tissue<br>GTGCCTCATCGTTTTGT<br>seq G T G T T T T G T T T T G T<br>Seq G T G T T T T T G T T T T G T<br>AMAGE AND A T T G T T T T G T   | Cancer tissue GTGCCTCATCGTTTTGGT EDM-seq 6 1 6 1 1 1 1 6 1 1 1 1 6 1 AMD-Seq 6 1 6 1 1 1 1 A 1 6 1 1 1 1 1 6 1   | Normal tissue<br>GTGCCTCATCGTTTTTGT<br>EDM-seq G T G T T T T T G T T T T T G T<br>AMD-Seq G T G T T T T T T G T T T T G T<br>AMD-Seq G T G T T T T T G T T T T G T   | Cancer tissue GTGCCTCATCGTTTTGT EDM-seqG1 GTTTTAT_GTTTTGT AMD-Seq G1 G1 TTTAT_GTTTTG T   |
| EDM-          | Normal tissue   | Cancer tissue  GTGCCTCATCGTTTTGT EDM-seq of of of the formation of the for | Normal tissue<br>EDM-seq of of it it is of it it  | Cancer tissue<br>GTGCCTCATCGTTTTGT<br>DM-seeg 61 61111 A 1 6 61111 6 1<br>AMD-Seeg 616111 A 1 6 61111 6 1<br>AMD-Seeg 616111 A 1 6 6 1111 6 1<br>AMD-Seeg 616111 A 1 6 6 1111 6 1<br>AMD-Seeg 61611 A 1 6 6 111 A 1 6 6 111 6 1<br>AMD-Seeg 61611 A 1 6 6 111 A 1 6 6 111 6 1<br>AMD-Seeg 61611 A 1 6 6 111 A 1 6 6 111 A 1 6 6 11 1 6 6 11 6 1 6  |
| EDM-          | Normal tissue   | Cancer tissue  COMPARENT OF CONTROLOGY CONTR | Normal tissue   EDM-seq of of it it is it is of it it it is of it it it is of it it is of it | Cancer tissue CDD-seed G G G G T T T A T G G T T T T G G T T T T   |
| EDM-<br>AMD-S | Normal tissue   GGGCCTCATCGTTTTGG   eq of of of the transformed of the transform   | Cancer tissue Concerting Concerti | Normal tissue   EDM-seq or   | Cancer tissue<br>CTCCCCTCCTCTTCTCT<br>CDM-seq 6 1 6 1 1 1 4 1 6 1 1 1 1 6 1<br>AMD-Seq 6 1 6 1 1 1 4 1 6 1 1 1 1 6 1<br>AMD-Seq 6 1 6 1 1 1 4 1 6 1 1 1 1 6 1<br>AMD-Seq 6 1 6 1 1 1 1 6 1 1 1 1 6 1 1 1 1 6 1 1 1 1 6 1 1 1 1 6 1 1 1 1 6 1 1 1 1 1 6 1 1 1 1 1 6 1 1 1 1 1 6 1 1 1 1 1 6 1 1 1 1 1 6 1 1 1 1 1 6 1 1 1 1 1 6 1 1 1 1 1 6 1 1 1 1 1 6 1 1 1 1 1 6 1 1 1 1 1 6 1 1 1 1 1 6 1 1 1 1 1 6 1 1 1 1 1 6 1   |
| EDM-          | Normal tissue     CONTRACTOR     Seq 0 TO   | Cancer tissue  | Normal lissue   EDM-seq or   | Cancer tissue COM-seq of of of the test of the test of |

## References

- 1. F. Tang, S. Liu, Q. Y. Li, J. Yuan, L. Li, Y. Wang, B. F. Yuan and Y. Q. Feng, *Chem Sci*, 2019, **10**, 4272–4281.
- 2. E. K. Schutsky, J. E. DeNizio, P. Hu, M. Y. Liu, C. S. Nabel, E. B. Fabyanic, Y. Hwang, F. D. Bushman, H. Wu and R. M. Kohli, *Nat Biotechnol*, 2018, **36**, 1083-1090.