1268–1277 *Nucleic Acids Research, 2019, Vol. 47, No. 3* *Published online 4 December 2018*
doi: 10.1093/nar/gky1218

N^6-Hydroxymethyladenine: a hydroxylation derivative of N^6-methyladenine in genomic DNA of mammals

Jun Xiong†, Tian-Tian Ye†, Cheng-Jie Ma, Qing-Yun Cheng, Bi-Feng Yuan* and Yu-Qi Feng

Key Laboratory of Analytical Chemistry for Biology and Medicine (Ministry of Education), Department of Chemistry, Wuhan University, Wuhan 430072, P.R. China

Received September 14, 2018; Revised November 04, 2018; Editorial Decision November 21, 2018; Accepted November 22, 2018

ABSTRACT

In addition to DNA cytosine methylation (5-methyl-2′-deoxycytidine, m^5dC), DNA adenine methylation (N^6-methyl-2′-deoxyadenosine, m^6dA) is another DNA modification that has been discovered in eukaryotes. Recent studies demonstrated that the content and distribution of m^6dA in genomic DNA of vertebrates and mammals exhibit dynamic regulation, indicating m^6dA may function as a potential epigenetic mark in DNA of eukaryotes besides m^5dC. Whether m^6dA undergoes the further oxidation in a similar way to m^5dC remains elusive. Here, we reported the existence of a new DNA modification, N^6-hydroxymethyl-2′-deoxyadenosine (hm^6dA), in genomic DNA of mammalian cells and tissues. We found that hm^6dA can be formed from the hydroxylation of m^6dA by the Fe^{2+}- and 2-oxoglutarate-dependent ALKBH1 protein in genomic DNA of mammals. In addition, the content of hm^6dA exhibited significant increase in lung carcinoma tissues. The increased expression of *ALKBH1* in lung carcinoma tissues may contribute to the increase of hm^6dA in DNA. Taken together, our study reported the existence and formation of hm^6dA in genomic DNA of mammals.

INTRODUCTION

DNA of living organisms is composed of four canonical nucleobases, adenine, guanine, cytosine, and thymine. In addition to these canonical bases, dynamic modifications that expand the genetic code are discovered to exist in DNA of prokaryotes and eukaryotes (1,2). DNA cytosine methylation (5-methyl-2′-deoxycytidine, m^5dC) is the best-characterized epigenetic modification in mammals (3). m^5dC now is viewed as the fifth nucleoside of DNA and participates in genomic imprinting, X-chromosome inactivation, and regulation of gene expression (4–6). Recently, 5-hydroxymethyl-2′-deoxycytidine (hm^5dC) was found as the sixth nucleoside of genomic DNA with potential roles in regulation of gene expression (7–10). 5-Formyl-2′-deoxycytidine (f^5dC) (11) and 5-carboxy-2′-deoxycytidine (ca^5dC) (12) were discovered in DNA in 2011. f^5dC and ca^5dC are commonly viewed as the intermediates in active DNA demethylation process, and they also play important roles in DNA replication, transcription and cell homeostasis (13–18).

In addition to DNA cytosine methylation, DNA adenine methylation (N^6-methyl-2′-deoxyadenosine, m^6dA) is also a naturally occurring DNA modification preserved in prokaryotes to eukaryotes (19). m^6dA was initially reported to be part of restriction-modification systems in prokaryotes (20). m^6dA has also been found to exert functional roles in DNA mismatch repair and gene regulatory processes in *Escherichia coli* (21). In 2015, three groups reported the existence and genome-wide distribution of m^6dA in different eukaryotes, including *Chlamydomonas reinhardtii* (22), *Caenorhabditis elegans* (23) and *Drosophila melanogaster* (24). More recently, m^6dA was also found in the genomes of vertebrates (25) and mammals (26–29). Although Schiffers et al. (30) reported no detectable m^6dA in genomic DNA of mouse embryonic stem cells and tissues, very recent genome-wide mapping analysis showed substantial m^6dA existing in human genome (29) and *Arabidopsis thaliana* genome (31). These studies demonstrated that m^6dA exhibited potential regulatory roles in gene activation, indicating m^6dA may serve as an additional epigenetic mark in DNA besides DNA cytosine methylation (32).

Interestingly, the content and distribution of m^6dA in genomic DNA altered during early embryogenesis of *D. melanogaster* (24), vertebrates and mammals (26), revealing that m^6dA may undergo dynamic regulation. As one homolog in the AlkB family proteins, fat mass and obesity-associated protein (FTO) was reported to mediate the oxidation of N^6-methyladenosine (m^6A) in RNA to form N^6-hydroxymethyladenosine (hm^6A) that is unstable and can lose a HCHO to form adenosine (33,34). In addition, active demethylation of m^5dC is achieved by 10–11 translocation (TET) proteins through forming the hydroxylation product of hm^5dC. Recent studies also showed that the proteins of the AlkB family, such as NMAD-1 in *C. elegans*

*To whom correspondence should be addressed. Tel: +86 27 68755595; Fax: +86 27 68755595; Email: bfyuan@whu.edu.cn
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

© The Author(s) 2018. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

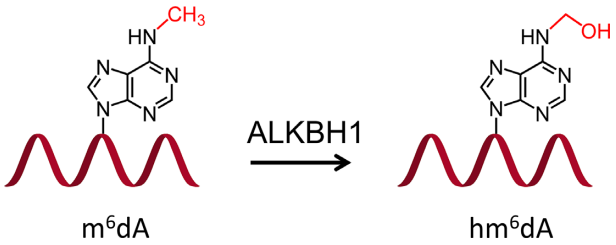

Figure 1. Schematic illustration of the formation of hm^{6}dA in DNA from the hydroxylation of m^{6}dA by the Fe^{2+}- and 2-oxoglutarate-dependent ALKBH1 protein.

(23), DMAD in *D. melanogaster* (24), and ALKBH1 in mouse and human (27,35), can regulate m^{6}dA level in genomic DNA. The *in-vitro* assay showed that ALKBH1 can induce the decline of m^{6}dA in DNA (36,37). Along this line, we hypothesize that *N*^{6}-hydroxymethyl-2′-deoxyadenosine (hm^{6}dA) can be formed from the hydroxylation of m^{6}dA by ALKBH1 in genomic DNA of mammals (Figure 1).

In this study, we reported the existence of hm^{6}dA in genomic DNA of mammalian cells and tissues. We demonstrated that ALKBH1 was able to convert m^{6}dA to hm^{6}dA *in vitro* and *in vivo*. In addition, the content of m^{6}dA exhibited significant decrease and hm^{6}dA exhibited significant increase companying with the increased expression of *ALKBH1* in lung carcinoma tissues compared to tumor-adjacent normal tissues.

MATERIALS AND METHODS

Chemicals and reagents

2′-Deoxycytidine (dC), 2′-deoxyguanosine (dG), 2′-deoxyadenosine (dA), thymidine (T), phosphodiesterase I, 20% formaldehyde solution and 20% D_2-formaldehyde solution were purchased from Sigma-Aldrich (St. Louis, MO, USA). *N*^{6}-methyl-2′-deoxyadenosine (m^{6}dA) and 5-methyl-2′-deoxycytosine (m^{5}dC) were purchased from Berry & Associates (Dexter, MI, USA). DNase I and alkaline phosphatase were from Takara Biotechnology Co. Ltd. (Dalian, China). Chromatographic grade methanol was purchased from Tedia Co. Inc. (Ohio, USA). All other solvents and chemicals were of analytical grade.

Synthesis of hm^{6}dA and D_2-hm^{6}dA

hm^{6}dA and D_2-hm^{6}dA were synthesized according to the previously reported method (38). Briefly, 10 mM of dA was mixed with 100 mM of formaldehyde or D_2-formaldehyde in phosphate buffer (pH 7.2) and incubated at 37°C for 12 h. hm^{6}dA or D_2-hm^{6}dA product was purified from the reaction mixture by HPLC using a Hisep C18-T column (250 mm × 4.6 mm i.d., 5 μm, Weltech Co., Ltd., Wuhan, China) with 0.01% formic acid (solvent A) and methanol (solvent B) as mobile phases. A gradient of 5% B for 3 min and 5–70% B for 20 min was used with a flow rate of 0.2 ml/min. hm^{6}dA or D_2-hm^{6}dA was collected and the concentration was determined by UV spectrophotometer (Metash Instruments Co., Ltd., Shanghai, China).

Evaluation of the stability of hm^{6}dA

The stability of hm^{6}dA was evaluated under different temperatures, pH values and times. As for the evaluation of the stability of hm^{6}dA under different temperatures, hm^{6}dA was incubated ranging from 4°C to 65°C for 2 h, or at 95°C for 5 min in enzymatic buffer (500 mM Tris–HCl, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM $ZnSO_4$, pH 7.0). As for the evaluation of the stability of hm^{6}dA under different pH values, hm^{6}dA was incubated in enzymatic buffer, pH ranging from 3.0 to 10.0 at 25°C for 2 h. As for the evaluation of the stability of hm^{6}dA under different time intervals, hm^{6}dA was incubated ranging from 30 min to 4 h at 25°C in enzymatic buffer (pH 7.0). The levels of hm^{6}dA under different conditions were measured by LC–MS with m^{5}dC as internal standard (IS).

Removal of potential contamination from bacterial genomic DNA by *Dpn* I digestion and size-exclusion ultrafiltration

The previously established *Dpn* I digestion combined with size-exclusion ultrafiltration method was used to remove the potential bacterial DNA contamination with slight modification (28). By virtue of selective cleavage of G(m^{6}dA)TC sites in DNA by *Dpn* I restriction enzyme (39,40), bacterial DNA can be cut into short fragments since adenines in bacterial DNA are heavily methylated to m^{6}dA. However, GATC sites in DNA cannot be cut by *Dpn* I. Thus, the DNA of mammals remains in large-size after digestion by *Dpn* I since m^{6}dA generally has very low abundance in genomic DNA of mammals. The resulting small DNA fragments from bacterial after *Dpn* I digestion can be removed from large-size genomic DNA of mammals using size-exclusion ultrafiltration. Briefly, a 50-μl mixture containing 10 μg of genomic DNA, 5 μl 10× cutsmart buffer (New England Biolabs), 80 units of *Dpn* I restriction enzyme (New England Biolabs) was incubated at 37°C for 1 h. Then the reaction products were transferred to a 100 kDa cut-off centrifugal filter (Millipore) and centrifuged at 8000 g for 3 min at 4°C. The centrifugal filter was washed three times with 500 μl of cold water for each time. The DNA retained on the filter was then collected and dissolved in water.

Cell culture and tissue samples

Human cervical carcinoma (HeLa) cells, human embryonic kidney (HEK293T) cells, GW5100 *E. coli* cells were obtained from the China Center for Type Culture Collection. HeLa and HEK293T cells were maintained in DMEM medium (Gibco) at 37°C under 5% CO_2 atmosphere. The medium was supplemented with 10% fetal bovine serum, 100 U/ml penicillin, and 100 μg/ml streptomycin (Gibco).

Male Sprague-Dawley rats (4 weeks old) were obtained from the Center for Animal Experiment/ABSL-3 Laboratory of Wuhan University and sacrificed to collect tissues which were stored under –80°C. The animal experiments were conducted in accordance to the guidelines described by the Animal Care and Ethics Committee of Wuhan University. A total of 10 pairs of human non-small cell lung carcinoma tissues and matched tumor-adjacent normal tissues without preoperative target therapy/chemotherapy from 10

patients were collected from Hubei Cancer Hospital. The lung carcinoma tissues and matched tumor-adjacent normal tissues were kept at –80°C. An approval for the study was granted by the Hubei Cancer Hospital Ethics Committee and met the declaration of Helsinki. All the experiments were performed in accordance with Hubei Cancer Hospital Ethics Committee's guidelines and regulations.

Genomic DNA extraction and enzymatic digestion

The genomic DNA from HeLa cells, HEK293T cells, *E. coli* cells, rat tissues and human lung carcinoma were extracted using DNAiso Reagent (Takara Biotechnology, Dalian, China) according to the manufacture's recommended procedure. All the DNA extraction was performed at 4°C. The concentration of the purified genomic DNA was determined on a B-500 spectrophotometer.

The extracted genomic DNA was first digested by *Dpn* I restriction enzyme followed by ultrafiltration to remove the potential contamination of bacterial DNA. The resulting genomic DNA was enzymatically digested under neutral conditions. Briefly, a 30-μl mixture containing 10 μg of DNA, 4 units of DNase I, 0.004 units of phosphodiesterase I, 30 units of alkaline phosphatase and 3 μl of enzymatic buffer (500 mM Tris–HCl, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM $ZnSO_4$, pH 7.0) was incubated at 25°C for 2 h. After adding 270 μl H_2O, the resulting solution was extracted with chloroform twice. The resulting aqueous layer was collected and lyophilized to dryness and then reconstituted in water followed by analysis with LC–MS.

Determination of hm^{6}dA by LC–MS analysis

LC–MS analysis of nucleosides was performed on the LC-ESI-MS/MS system consisting of a Shimadzu LC-20AD HPLC (Tokyo, Japan) and an AB 3200 QTRAP mass spectrometer (Applied Biosystems, Foster City, CA) with an electrospray ionization source (Turbo Ionspray). Data acquisition and processing were performed using AB SCIEX Analyst 1.5 Software (Applied Biosystems, Foster City, CA, USA). The LC separation was performed on a Shimadzu VP-ODS column (250 mm × 2.1 mm i.d., 5 μm, Tokyo, Japan) with a flow rate of 0.2 ml/min at 35°C. 2 mM NH_4HCO_3 in water (solvent A) and methanol (solvent B) were employed as mobile phase. A gradient of 0–5 min 5% B, 5–15 min 5–25% B, 15–28 min 25–70% B and 30–40 min 5% B was used.

The mass spectrometry detection was performed under positive electrospray ionization (ESI) mode. The nucleosides were monitored by multiple reaction monitoring (MRM) mode. The MRM parameters of all nucleosides were optimized to achieve maximal detection sensitivity. The mass transitions of nucleosides are listed in Table S1 in Supporting Information. Quantification of measured nucleosides was carried out according to previously described method (28) and detailed information can be found in the Supporting Information.

High-resolution mass spectrometry analysis

The hm^{6}dA was examined by LTQ-Orbitrap Elite high-resolution mass spectrometer (Thermo-Fisher Scientific, Waltham, MA, USA) equipped with an ESI source and Dionex ultimate 3000 UPLC system (Thermo-Fisher Scientific, Waltham, MA, USA). The LC separation conditions were the same as that on AB 3200 QTRAP mass spectrometer system. Full MS scans were acquired under positive-ion mode at a resolution of 60,000. The molecular mass of the hm^{6}dA was listed as the precursor ion for MS^2 analysis. Collision induced dissociation (CID) with the collision energy of 35 eV was used. The fragments were acquired with a mass range of m/z 100–500 at a resolution of 60 000 and an acquisition time of 10 ms. The source and ion transfer parameters applied were as follows: heater temperature, 300°C; capillary temperature, 350°C; sheath gas flow, 35 arbitrary units; auxiliary gas flow, 7 arbitrary units; spray voltage, 3.5 kV; capillary voltage, 35 V; S-lens RF level, 60%. The data analysis was achieved using Xcalibur v3.0 (Thermo-Fisher Scientific, Waltham, MA, USA).

Expression and purification of recombinant human ALKBH1 protein

To express the recombinant human ALKBH1protein in *E. coli*, the pGEX-ALKBH1 plasmid was constructed by inserting the full length coding sequence of human *ALKBH1* into the vector of pGEX-4T1, which contains the glutathione S-transferase (GST) tag and *Eco*R I/*Xho* I cloning sites (GE Healthcare, USA). The forward and reverse primers used for PCR amplification of *ALKBH1* gene were 5′-CCGGAATTCATGG GGAAGATGGCAGCGG-3′ and 5′-CCGCTCGAGTC AGCTGTCAGGGTTTATCC-3′, respectively. The site-specific mutation of ALKBH1 (D233A) was generated by overlap extension PCR with a pair of overlap PCR primers (forward primer: 5′-ACACTGGGAATCC ACGTAGCCAGATCTGAGCTAGATCACT-3′; reverse primer: 5′-AGTGATCTAGCTCAGATCTGGCTACGT GGATTCCCAGTGT-3′) (41). The constructs were confirmed by DNA sequencing and expressed in BL21 (codon plus) strain. Protein expression was induced using 0.5 mM IPTG (isopropyl-β-D-thiogalactopyranoside) for 12 h at 18°C. Recombinant ALKBH1 and ALKBH1 (D233A) proteins were purified with GE Healthcare Life Sciences™ Glutathione Sepharose™ 4B Media following the manufacturer's protocol.

Biochemical assay of ALKBH1-mediated formation of hm^{6}dA *in vitro*

The assay of ALKBH1-mediated formation of hm^{6}dA was performed in a 50-μl reaction mixture containing 5 pmol of DNA carrying m^{6}dA, 2 μg of recombinant ALKBH1 (or D233A mutant) protein, 50 μM KCl, 1 mM $MgCl_2$, 50 μM HEPES (pH 7.0), 2 mM ascorbic acid, 1 mM 2-oxoglutarate (2-KG), and 1 mM $(NH_4)_2Fe(SO_4)_2{\cdot}6H_2O$. The reaction was performed at 37°C for 1 h. Then the DNA in the reaction mixture was purified using QIAquick Nucleotide Removal Kit (QIAGEN, USA) to remove salts and protein according to the manufacture's recommended procedure. The purified DNA was digested to nucleosides and subjected to LC–MS analysis. The sequences of different DNA used in this assay are listed in Table S2 in Supporting Information.

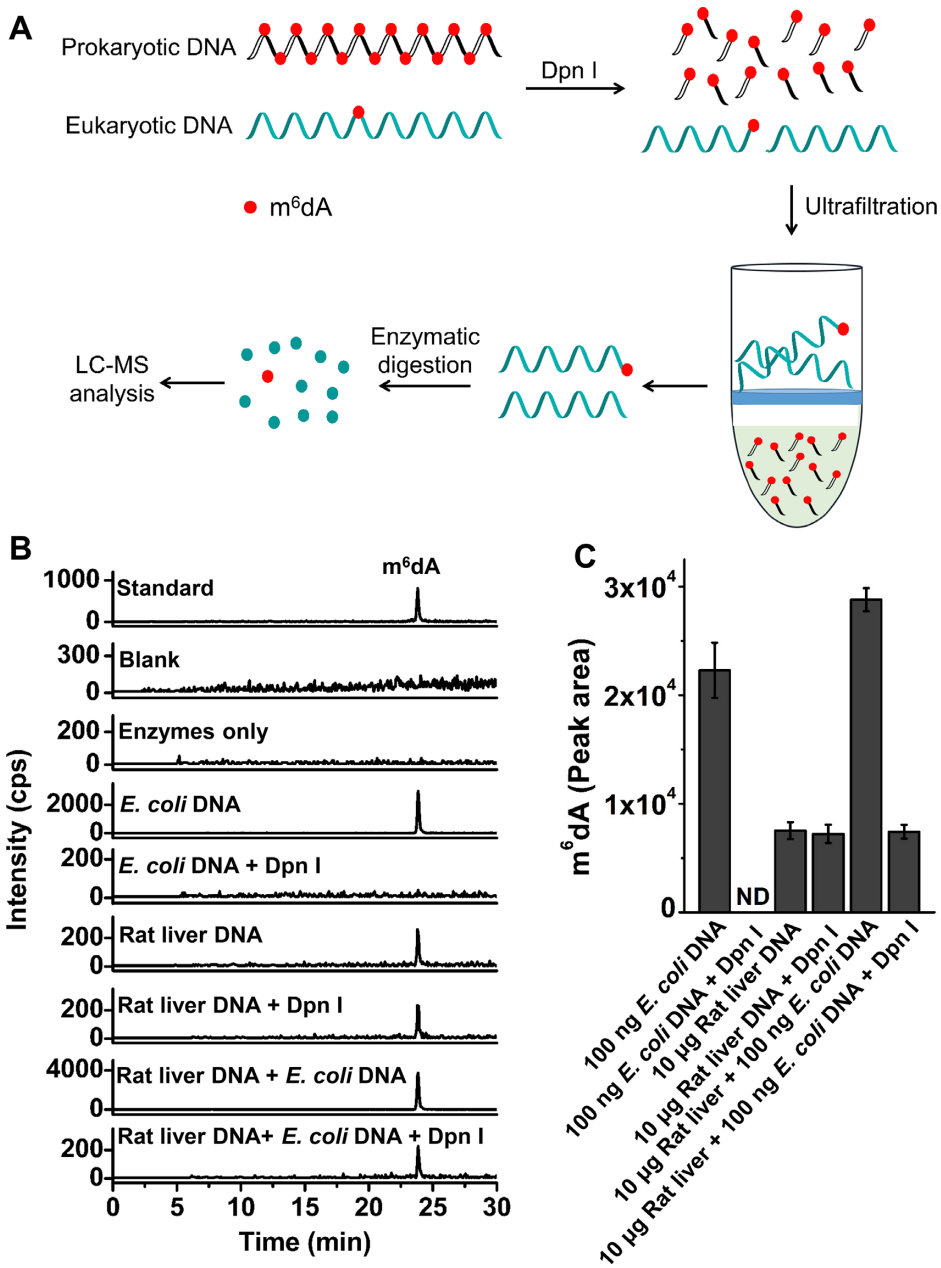

Figure 2. Evaluation of the removal of potential bacterial DNA contamination from genomic DNA of mammals by *Dpn* I digestion coupled with size-exclusion ultrafiltration. (**A**) Schematic illustration of *Dpn* I digestion coupled with size-exclusion ultrafiltration method. Bacterial DNA can be cut into short fragments at G(m^{6}dA)TC sites since adenines in bacterial DNA are heavily methylated to m^{6}dA. The DNA of mammals remains in large-size after digestion by *Dpn* I since m^{6}dA generally has very low abundance in genomic DNA of mammals. The resulting small DNA fragments from bacterial after *Dpn* I digestion can be removed from large-size genomic DNA of mammals using size-exclusion ultrafiltration. (**B**) Extracted-ion chromatograms for the detection of m^{6}dA from different samples by LC–MS analysis. (**C**) Measured contents of m^{6}dA from different samples ($n = 3$).

Overexpression and knockdown of ALKBH1 by transient transfection

The plasmid of pCMV-ALKBH1 was constructed by fusing the full-length cDNA of *ALKBH1* into the vector of pCMV-EGFP at the *Xho* I/*Kpn* I cloning sites. The forward and reverse primers for PCR amplification of *ALKBH1* gene were 5′-CCGCTCGAGATGGGGAAGATGGCAG CGG-3′ and 5′-CGGGGTACCTCAGCTGTCAGGGT TTATCC-3′, respectively.

RNA knockdown of *ALKBH1* was performed using siRNA (Takara Biotechnology) against human *ALKBH1* mRNA. The non-targeting siRNA was used as a negative control. The sequences of *ALKBH1* siRNA and control siRNA are 5′-GGAUGACCAGAAUAGCGAA TT-3′/3′-TTCCUACUGGUCUUAUCGCUU-5′ and 5′-UUCUCCGAACGUGUCACGUTT-3′/3′-TTAAG AGGCUUGCACAGUGCA-5′, respectively.

Human HEK293T cells were transfected with pCMV-ALKBH1 plasmid or siRNA using Lipofectamine 3000 (In-

vitrogen, USA) according to manufacturer's instruction. Transfection of HEK293T cells with the vector of pCMV-EGFP or control siRNA was used as the control for overexpression and RNA knockdown, respectively. The culture medium was replaced at 10 h after transfection. HEK293T cells were harvested and genomic DNA was isolated at 48 h after transfection.

Quantitative real-time PCR and western blot

For the relative quantification of the expression of *ALKBH1*, 1 μg of isolated total RNA was used to generate cDNA using PrimeScript™ RT reagent Kit with gDNA Eraser (Takara Biotechnology). Quantitative PCR (qPCR) was performed using a CFX96™ Real-Time PCR Detection System (Bio-Rad Laboratories) and SYBR® Premix Ex Taq II (Tli RNaseH Plus) (Takara Biotechnology) according to the manufacturer's instructions. The levels of gene expression were normalized to glyceraldehyde 6-phosphate dehydrogenase gene (*GAPDH*). PCR primers (Sangon Biotech, Shanghai, China) sequences are: *GAPDH* forward primer 5′-GCCAGCCTCGTCCCGTAGACA-3′, *GAPDH* reverse primer 5′-CAACAATCTCCACTTTGCC ACTGC-3′; *ALKBH1* forward primer 5′-CACCATTCT GCTGTGCCCTA-3′, *ALKBH1* reverse primer 5′-CAA GCTGCCTACCCTCAGAC-3′.

The overexpression and siRNA knockdown of *ALKBH1* were also examined by western blot. Briefly, human HEK293T cell nuclear lysates were prepared using Nucleoprotein Extraction Kit (Sangon, Shanghai, China). Protein concentration was measured by the BCA (bicinchoninic acid) assay according to the manufacture's recommended procedure (Beyotime Biotech Inc, Shanghai, China). Cell nuclear were subjected to Western blot analysis. Antibodies that specifically recognized ALKBH1 (Abcam, ab12889, Cambridge, MA) and GAPDH (Abcam, ab181602, Cambridge, MA) were used at 1:2000 and 1:8000 dilutions, respectively. Horseradish peroxidase-conjugated secondary goat anti-rabbit antibody (Abcam, ab6721, Cambridge, MA) was used at a 1:10 000 dilution.

Statistical analysis

The experimental data were processed and analysed with SPSS 19.0 software (IBM SPSS Inc, USA). The ANOVA (Analysis of Variance) test was performed to evaluate the differences of m^6dA modification between multiple conditions of *in-vitro* biochemical assay. The paired *t*-test was performed to evaluate the differences of m^6dA and hm^6dA between lung carcinoma tissues and tumor-adjacent normal tissues. And P value < 0.05 was considered to have statistical significance.

RESULTS AND DISCUSSION

Evaluation of the stability of hm^6dA

The discovery of m^6dA in genomic DNA of eukaryotes have spurred the study in unravelling the functions of m^6dA. In this study, we investigated the existence of hm^6dA in DNA of mammals and the conversion of m^6dA to hm^6dA by ALKBH1 *in vitro* and *in vivo*. To this end, we first synthesized hm^6dA and D_2-hm^6dA standards and the LC chromatograms for the purification were shown in Figure S1 in Supporting Information. The mass spectrometry analysis demonstrated that hm^6dA and D_2-hm^6dA compounds were successfully synthesized (Figure S2 in Supporting Information). Then we evaluated the stability of hm^6dA and the results showed that hm^6dA was less stable at high temperature (Figure S3A in Supporting Information) and high pH (Figure S3B in Supporting Information). hm^6dA can be converted to dA by losing a HCHO. We found hm^6dA was stable at least for 4 h under neutral condition (pH 7.0) at 25°C (Figure S3C in Supporting Information). Thus, to determine hm^6dA, the DNA extraction and handling was processed at 4°C, and enzymatic digestion of genomic DNA was performed under neutral condition (pH 7.0) at 25°C for 2 h.

Determination of hm^6dA in genomic DNA of mammalian cells

Bacterial DNA contains abundant m^6dA and may cause potential contamination for the analysis of m^6dA and hm^6dA from mammalian genomic DNA. Here, we used the previously established *Dpn* I digestion combined with size-exclusion ultrafiltration method (28) to remove the potential bacterial DNA (Figure 2A). In addition, we also examined potential contamination of m^6dA from enzymes used in DNA digestion. In this regard, we analyzed the sample with only adding enzymes and omitting the genomic DNA. The results showed that no m^6dA was observed from the enzymes, suggesting that these enzymes didn't contain m^6dA (Figure 2B).

Then we used *E. coli* DNA and rat liver DNA to evaluate the performance of the method on removing bacterial DNA. The peak areas of m^6dA in 100 ng of *E. coli* DNA and 10 μg of rat liver DNA were $\sim 2.2 \times 10^4$ and 0.75×10^4, respectively, by direct analysis without *Dpn* I digestion and size-exclusion ultrafiltration (Figure 2C), suggesting that the content of m^6dA in 10 μg of rat liver DNA was less than that in 100 ng of *E. coli* DNA. Therefore, we used 100 ng bacterial DNA, and 10 μg of rat liver DNA spiked with 100 ng *E. coli* DNA to perform the evaluation. The results showed that no m^6dA was detected from 100 ng of *E. coli* DNA after *Dpn* I digestion and size-exclusion ultrafiltration, suggesting the presence of *E. coli* DNA is below the detection limit (Figure 2C). The peak area of m^6dA in 10 μg of rat liver DNA slightly decreased after *Dpn* I digestion and size-exclusion ultrafiltration compared to the direct analysis (Figure 2C). However, after *Dpn* I digestion and size-exclusion ultrafiltration, the peak area of m^6dA in 10 μg of rat liver DNA spiked with 100 ng *E. coli* DNA dramatically decreased to the level of 10 μg of rat liver DNA (Figure 2C). These results suggested that rat liver DNA contained almost no bacterial DNA, and bacterial DNA contamination can also be efficiently removed if it existed in rat liver DNA.

Using this method, we next explored the existence of hm^6dA in mammalian cells and tissues. Genomic DNA was pretreated by *Dpn* I digestion and size-exclusion ultrafiltration, and DNA enzymatic digestion were performed un-

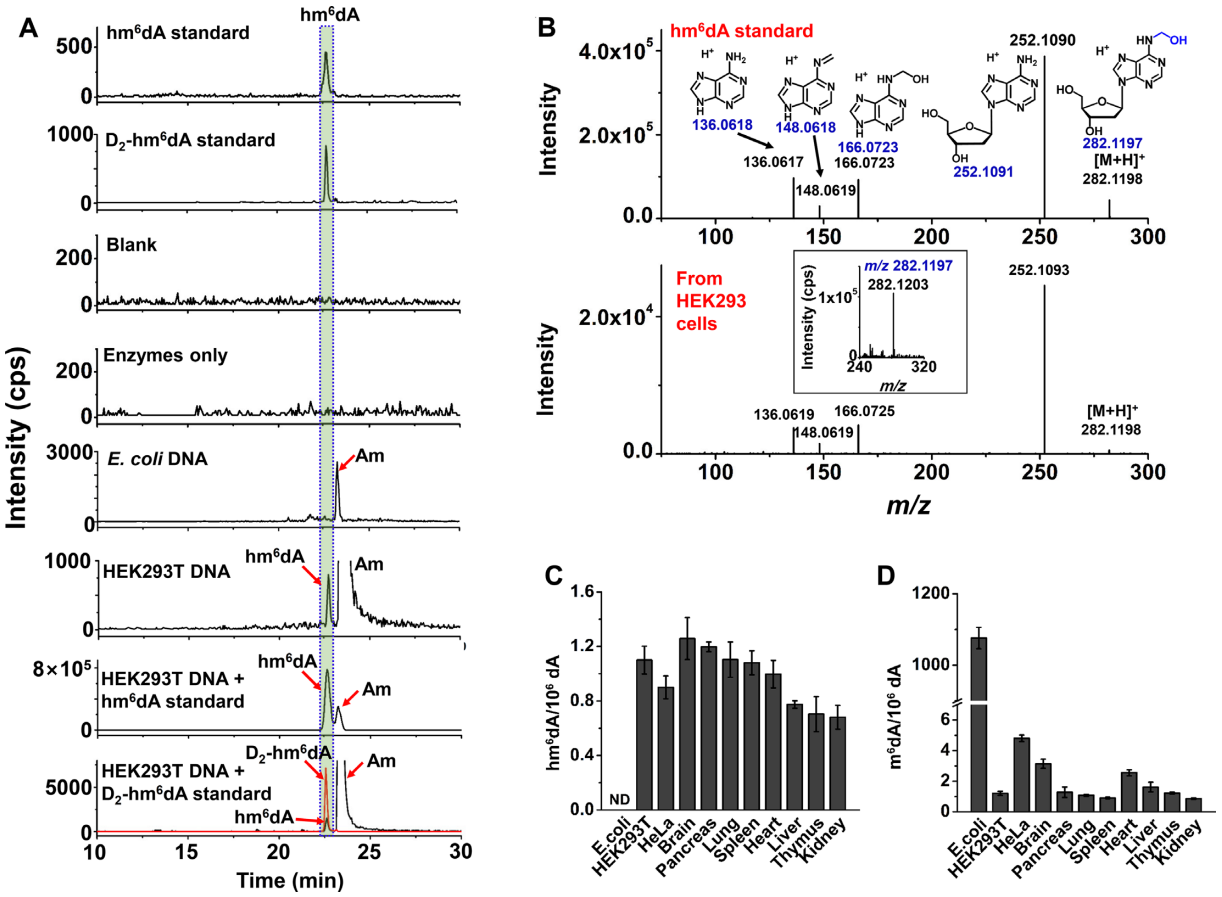

Figure 3. Determination of hm^{6}dA in mammalian cells and tissues. (**A**) Extracted-ion chromatograms for the detection of hm^{6}dA from different samples. The synthesized hm^{6}dA or D_2-hm^{6}dA standard was added into the enzymatically digested DNA of HEK293T cells to confirm the existence of hm^{6}dA in HEK293T cells. Am, 2′-*O*-methyladenosine. (**B**) High-resolution mass spectrometry analysis of hm^{6}dA standard (upper spectrum) and hm^{6}dA from genomic DNA of HEK293T cells (bottom spectrum). (**C**) Measured amounts of hm^{6}dA in various samples ($n = 3$). (**D**) Measured amounts of m^{6}dA in various samples ($n = 3$).

der neutral condition (pH 7.0) at 25°C for 2 h. The retention time (t_R = 22.6 min) of the compound detected in genomic DNA of HEK293T cells was similar to the hm^{6}dA and D_2-hm^{6}dA standards by LC–MS analysis under MRM detection mode (282.2 → 136.1 for hm^{6}dA; 284.2 → 136.1 for D_2-hm^{6}dA) (Figure 3A), indicating that the compound should be hm^{6}dA. In addition, the peak intensity increased while the synthesized hm^{6}dA or D_2-hm^{6}dA standard was added into the enzymatically digested DNA of HEK293T cells (Figure 3A), further supporting the existence of hm^{6}dA in HEK293T cells. On the contrary, hm^{6}dA was undetectable in *E. coli* DNA or in the sample with only adding enzymes and omitting the genomic DNA (Figure 3A), excluding the possibility that the detected hm^{6}dA from *E. coli* or enzymes used in DNA digestion. The overall chromatograms for the detection of canonical nucleosides (dC, dG, T and dA) as well as the m^{6}dA and hm^{6}dA are shown in Figure S4 in Supporting Information.

Moreover, the high-resolution mass spectrometry analysis showed the product-ion spectrum of detected compound from HEK293T cells was identical to the authentic hm^{6}dA standard (Figure 3B), further confirming the detected compound was hm^{6}dA. The quantification results showed that the content of hm^{6}dA ranged from 0.7 to 1.3 modifications per 10^6 dA in HEK293T cells, HeLa cells and rat tissues, which was within the same order of magnitude as that of m^{6}dA (Figure 3C and D). It should be noted that the isolated DNA may contain small amount of DNA from mitochondria, which may also contribute to the measured m^{6}dA and hm^{6}dA.

ALKBH1 oxidizes m^{6}dA to form hm^{6}dA

With the detected hm^{6}dA from mammalian cells, we next assessed the capability of recombinant human ALKBH1 protein in inducing the oxidation of m^{6}dA in DNA by conducting an *in-vitro* reaction with the use of a single-stranded DNA carrying a single m^{6}dA (Figure 4). The SDS-PAGE confirmed the successful preparation of recombinant wild-type and mutant ALKBH1 proteins (Figure S5 in Supporting Information). We first examined the potential existence of m^{6}dA and hm^{6}dA from recombinant human ALKBH1 protein and mutant ALKBH1 protein (D233A). The results showed that no m^{6}dA and hm^{6}dA were detected from these recombinant proteins (Figure S6 in Supporting Information).

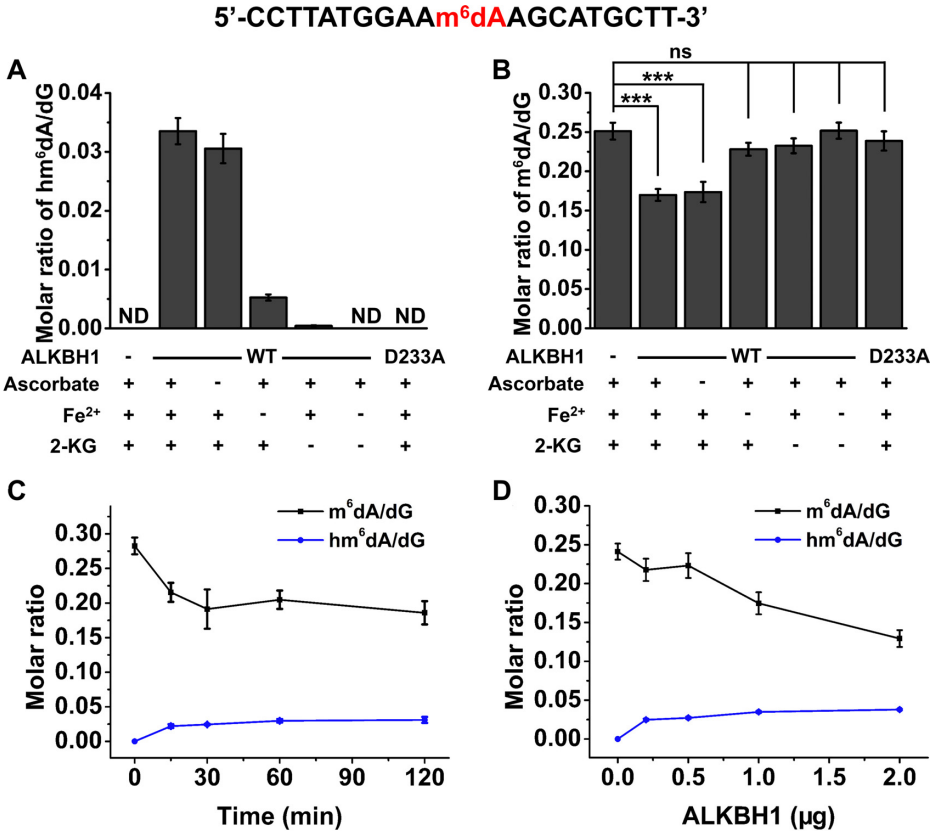

Figure 4. ALKBH1 oxidizes m^6dA to form hm^6dA *in vitro*. (**A**) Amounts of detected hm^6dA in the *in-vitro* reaction under different conditions. (**B**) Amounts of detected m^6dA in the *in-vitro* reaction under different conditions. (**C**) Time-dependent formation of hm^6dA in single-stranded DNA. (**D**) Amounts of hm^6dA and m^6dA with the increased amount of ALKBH1 protein. WT, wild-type ALKBH1 protein; D233A, mutant ALKBH1 protein (D233A); 2-KG, 2-oxoglutarate. ND, not detected. $^{**}P < 0.01$ ($n = 3$). NS, not significant. The ANOVA test was performed to evaluate the differences of m^6dA modification between multiple conditions of *in-vitro* biochemical assay.

LC–MS analysis of the nucleoside mixture from the enzymatic digestion of the DNA isolated from the reaction mixture revealed the formation of hm^6dA (Figure 4A), which is accompanied by a decrease in the level of m^6dA (Figure 4B). It is of note that omitting Fe^{2+} in the reaction buffer led to a decrease in the formation of hm^6dA, whereas exclusion of 2-oxoglutarate (2-KG) in the reaction buffer nearly abolished the ALKBH1-catalyzed formation of hm^6dA (Figure 4A). Both omission of Fe^{2+} and 2-KG resulted in no detectable formation of hm^6dA (Figure 4A). Omission of ascorbate showed no obvious effect on catalytic activity of ALKBH1 (Figure 4A). However, the oxidation activity of ALKBH1 was abolished by a point mutant at a critical residue (D233A), which renders the catalytic site ineffective (Figure 4A and B) (27). Together, the results supported that hm^6dA arose from the Fe^{2+}- and 2-KG-dependent dioxygenase activity of ALKBH1.

The level of m^6dA showed a time-dependent decrease and hm^6dA showed time-dependent increase by incubation with ALKBH1 protein (Figure 4C). In addition, with the increased amount of ALKBH1 protein, the levels of m^6dA and hm^6dA decreased and increased, respectively (Figure 4D). We also found that ALKBH1 displayed a higher activity toward single-stranded DNA (ssDNA1 listed in Table S2 in Supporting Information) than hemi-methylated double-stranded DNA (dsDNA listed in Table S2 in Supporting Information) in the same sequence context (Figures S7–S10 in Supporting Information). In addition, we further tested the oxidation activity of ALKBH1 with a longer m^6dA-containing single-stranded DNA (ssDNA2 listed in Table S2 in Supporting Information). The results demonstrated that ALKBH1 also can convert m^6dA to hm^6dA (Figures S7–S10 in Supporting Information). Taken together, these results suggested that ALKBH1 was capable to convert m^6dA to hm^6dA *in vitro*.

We next performed the overexpression and knockdown of *ALKBH1* gene in HEK293T cells. The overexpression and knockdown of *ALKBH1* gene were confirmed by quantitative real-time PCR analysis and western blot analysis (Figure S11 in Supporting Information). The LC–MS results demonstrated that overexpression of *ALKBH1* caused marked decline in the level of m^6dA, and increase in the level of hm^6dA ($P < 0.01$, Figure 5). On the contrary, the knockdown of *ALKBH1* by siRNA induced a significant increase of the level of m^6dA and decrease of the level of hm^6dA ($P < 0.01$, Figure 5). This study revealed ALKBH1 can convert m^6dA to hm^6dA *in vivo*. While, it's still possible there may exist other enzymes that can also function on the conversion of m^6dA to hm^6dA besides ALKBH1. We cannot exclude this possibility at current stage.

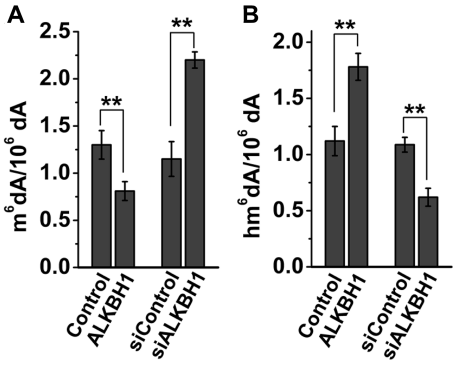

Figure 5. Overexpression and knockdown of *ALKBH1* induced the contents change of m^6dA and hm^6dA. (**A**) Measured m^6dA in HEK293T cells upon overexpression and knockdown of *ALKBH1*. (**B**) Measured hm^6dA in HEK293T cells upon overexpression and knockdown of *ALKBH1*. ** $P < 0.01$ ($n = 3$). Two-side unpaired *t*-test was performed.

It should be noted that hm^6dA has been previously observed as a product of DNA adduct induced by endogenous and exogenous formaldehyde (38,42). However, we failed to detect the other formaldehyde-induced DNA adducts, N^2-hydroxymethyl-2′-deoxyguanosine (hm^2dG) and N^2-hydroxymethyl-2′-deoxycytidine (hm^4dC) *in vivo* (Figure S12 in Supporting Information). Moreover, when we conducted the *in-vitro* reaction of ALKBH1 protein with a single-stranded or double-stranded DNA carrying no m^6dA (ssDNA control and dsDNA control listed in Table S2 in Supporting Information), we did not observe hm^6dA, hm^2dG and hm^4dC. The results suggested that enzymatic oxidation of m^6dA should be the major source of hm^6dA.

The level of hm^6dA increases in genomic DNA of lung carcinoma tissues

To further understand the physiological consequence of hm^6dA, we quantified the contents of m^6dA and hm^6dA in lung carcinoma tissues and tumor-adjacent normal tissues. A total of 20 tissues from 10 lung cancer patients were analyzed. The results showed the significant decrease of the level of m^6dA in lung carcinoma tissues compared to tumor-adjacent normal tissues (m^6dA, $P = 0.004$, Figure 6A and Figure S13A in Supporting Information), which is similar to the recent report that the level of m^6dA was declined in liver cancer tissues (29).

On the contrary, we observed the significant increase of the content of hm^6dA in lung carcinoma tissues compared to tumor-adjacent normal tissues (hm^6dA, $P = 0.001$, Figure 6B and Figure S13B in Supporting Information). To further investigate the potential reason that induced the decrease of m^6dA and increase of hm^6dA in DNA, we examined the expression of *ALKBH1* since it can convert m^6dA to hm^6dA. The result showed that the mRNA expression of *ALKBH1* were generally upregulated in lung carcinoma tissues compared to tumor-adjacent normal tissues (Figure 6E), which indicated the increased expression of *ALKBH1* in lung carcinoma tissues may contribute to the decrease of m^6dA and increase of hm^6dA in DNA. The quantification results also showed the significant decrease of 5-hmdC ($P < 0.001$) and no obvious change of m^5dC ($P = 0.059$) in lung

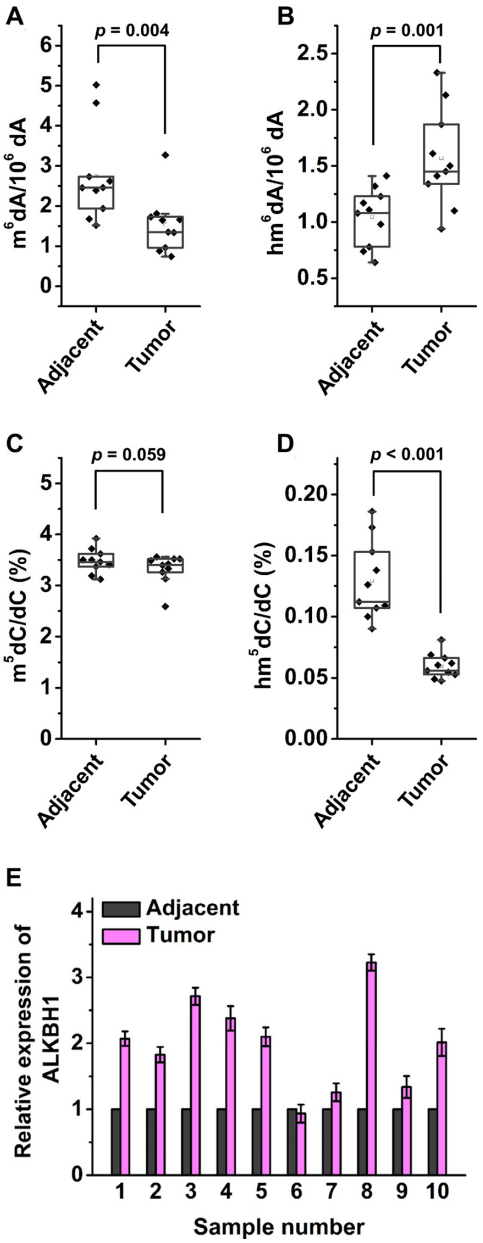

Figure 6. Quantification and statistical analysis of DNA modifications and the relative mRNA level of *ALKBH1* in human lung carcinoma tissues and tumor-adjacent normal tissues. Quantification and statistical analysis of m^6dA (**A**), hm^6dA (**B**), m^5dC (**C**) and hm^5dC (**D**) by LC–MS/MS. (**E**) The relative mRNA level of *ALKBH1*. *GAPDH* was used as endogenous control and triplicate measurements were carried out. A total of 20 tissues from 10 lung cancer patients were analysed. Two-side paired *t*-test was performed.

carcinoma tissues compared to tumor-adjacent normal tissues (Figure 6C and D, Supplementary Figures S13C and S13D), which is consistent with previous study (43).

These results demonstrated that the levels of both m^6dA and hm^6dA significantly changed in lung carcinoma tissues. However, the potential role of hm^6dA in shaping gene expression remains elusive. Although the level of hm^6dA in genomic DNA of mammals is low, hm^6dA may evolve specialized functions similar to hm^5dC which is also present in

low abundance in genomic DNA but with important epigenetic regulation roles. Future study on mapping hm^{6}dA in genomes should be helpful to further decipher its biological functions.

In summary, we reported the existence of hm^{6}dA in genomic DNA of mammalian cells and tissues by LC–MS analysis. We further showed that the Fe^{2+}- and 2-KG-dependent ALKBH1 can convert m^{6}dA in formation of hm^{6}dA *in vitro* and *in vivo*. In addition, the content of hm^{6}dA showed significant increase in lung carcinoma tissues compared to tumor-adjacent normal tissues. This study provided new insight for understanding the dynamic regulation of DNA adenine methylation in eukaryotes.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Chu-Bo Qi (Hubei Cancer Hospital, Wuhan, P.R. China) for providing the lung carcinoma tissues, and Prof. Yan Wu (Wuhan University, P.R. China) and Xiao-Ming Yin (Wuhan University, P.R. China) for the help on western blot experiments.

FUNDING

National Natural Science Foundation of China [21522507, 21672166, 21635006, 21721005, 21728802]. Funding for open access charge: National Natural Science Foundation of China.
Conflict of interest statement. None declared.

REFERENCES

1. Carell,T., Brandmayr,C., Hienzsch,A., Muller,M., Pearson,D., Reiter,V., Thoma,I., Thumbs,P. and Wagner,M. (2012) Structure and function of noncanonical nucleobases. *Angew. Chem. Int. Ed. Engl.*, **51**, 7110–7131.
2. Liu,T., Ma,C.J., Yuan,B.F. and Feng,Y.Q. (2018) Modificaomics: deciphering the functions of biomolecule modifications. *Sci. China Chem.*, **61**, 381–392.
3. Shen,H. and Laird,P.W. (2013) Interplay between the cancer genome and epigenome. *Cell*, **153**, 38–55.
4. Wu,H. and Zhang,Y. (2014) Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell*, **156**, 45–68.
5. Chen,Y., Hong,T., Wang,S., Mo,J., Tian,T. and Zhou,X. (2017) Epigenetic modification of nucleic acids: from basic studies to medical applications. *Chem. Soc. Rev.*, **46**, 2844–2872.
6. Liu,S. and Wang,Y. (2015) Mass spectrometry for the assessment of the occurrence and biological consequences of DNA adducts. *Chem. Soc. Rev.*, **44**, 7829–7854.
7. Munzel,M., Globisch,D. and Carell,T. (2011) 5-Hydroxymethylcytosine, the sixth base of the genome. *Angew. Chem Int. Ed. Engl.*, **50**, 6460–6468.
8. Kriaucionis,S. and Heintz,N. (2009) The nuclear DNA Base 5-Hydroxymethylcytosine is present in purkinje neurons and the brain. *Science*, **324**, 929–930.
9. Tahiliani,M., Koh,K.P., Shen,Y., Pastor,W.A., Bandukwala,H., Brudno,Y., Agarwal,S., Iyer,L.M., Liu,D.R., Aravind,L. *et al.* (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, **324**, 930–935.
10. Zhang,Y., Liu,L., Guo,S., Song,J., Zhu,C., Yue,Z., Wei,W. and Yi,C. (2017) Deciphering TAL effectors for 5-methylcytosine and 5-hydroxymethylcytosine recognition. *Nat. Commun.*, **8**, 901.
11. Ito,S., Shen,L., Dai,Q., Wu,S.C., Collins,L.B., Swenberg,J.A., He,C. and Zhang,Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, **333**, 1300–1303.
12. He,Y.F., Li,B.Z., Li,Z., Liu,P., Wang,Y., Tang,Q., Ding,J., Jia,Y., Chen,Z., Li,L. *et al.* (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, **333**, 1303–1307.
13. Spruijt,C.G., Gnerlich,F., Smits,A.H., Pfaffeneder,T., Jansen,P.W., Bauer,C., Munzel,M., Wagner,M., Muller,M., Khan,F. *et al.* (2013) Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell*, **152**, 1146–1159.
14. Su,M., Kirchner,A., Stazzoni,S., Muller,M., Wagner,M., Schroder,A. and Carell,T. (2016) 5-Formylcytosine could be a semipermanent base in specific genome sites. *Angew. Chem. Int. Ed. Engl.*, **55**, 11797–11800.
15. Kellinger,M.W., Song,C.X., Chong,J., Lu,X.Y., He,C. and Wang,D. (2012) 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.*, **19**, 831–833.
16. Xing,X.W., Liu,Y.L., Vargas,M., Wang,Y., Feng,Y.Q., Zhou,X. and Yuan,B.F. (2013) Mutagenic and cytotoxic properties of oxidation products of 5-methylcytosine revealed by next-generation sequencing. *PLoS One*, **8**, e72993.
17. Xia,B., Han,D., Lu,X., Sun,Z., Zhou,A., Yin,Q., Zeng,H., Liu,M., Jiang,X., Xie,W. *et al.* (2015) Bisulfite-free, base-resolution analysis of 5-formylcytosine at the genome scale. *Nat. Methods*, **12**, 1047–1050.
18. Zhu,C., Gao,Y., Guo,H., Xia,B., Song,J., Wu,X., Zeng,H., Kee,K., Tang,F. and Yi,C. (2017) Single-Cell 5-Formylcytosine landscapes of mammalian early embryos and ESCs at Single-Base resolution. *Cell Stem Cell*, **20**, 720–731.
19. Heyn,H. and Esteller,M. (2015) An adenine code for DNA: A second life for N6-Methyladenine. *Cell*, **161**, 710–713.
20. Luo,G.Z. and He,C. (2017) DNA N(6)-methyladenine in metazoans: functional epigenetic mark or bystander? *Nat. Struct. Mol. Biol.*, **24**, 503–506.
21. Fang,G., Munera,D., Friedman,D.I., Mandlik,A., Chao,M.C., Banerjee,O., Feng,Z., Losic,B., Mahajan,M.C., Jabado,O.J. *et al.* (2012) Genome-wide mapping of methylated adenine residues in pathogenic Escherichia coli using single-molecule real-time sequencing. *Nat. Biotechnol.*, **30**, 1232–1239.
22. Fu,Y., Luo,G.Z., Chen,K., Deng,X., Yu,M., Han,D., Hao,Z., Liu,J., Lu,X., Dore,L.C. *et al.* (2015) N(6)-methyldeoxyadenosine marks active transcription start sites in chlamydomonas. *Cell*, **161**, 879–892.
23. Greer,E.L., Blanco,M.A., Gu,L., Sendinc,E., Liu,J., Aristizabal-Corrales,D., Hsu,C.H., Aravind,L., He,C. and Shi,Y. (2015) DNA Methylation on N(6)-Adenine in C. elegans. *Cell*, **161**, 868–878.
24. Zhang,G., Huang,H., Liu,D., Cheng,Y., Liu,X., Zhang,W., Yin,R., Zhang,D., Zhang,P., Liu,J. *et al.* (2015) N(6)-methyladenine DNA modification in Drosophila. *Cell*, **161**, 893–906.
25. Koziol,M.J., Bradshaw,C.R., Allen,G.E., Costa,A.S., Frezza,C. and Gurdon,J.B. (2016) Identification of methylated deoxyadenosines in vertebrates reveals diversity in DNA modifications. *Nat. Struct. Mol. Biol.*, **23**, 24–30.
26. Liu,J., Zhu,Y., Luo,G.Z., Wang,X., Yue,Y., Wang,X., Zong,X., Chen,K., Yin,H., Fu,Y. *et al.* (2016) Abundant DNA 6mA methylation during early embryogenesis of zebrafish and pig. *Nat. Commun.*, **7**, 13052.
27. Wu,T.P., Wang,T., Seetin,M.G., Lai,Y., Zhu,S., Lin,K., Liu,Y., Byrum,S.D., Mackintosh,S.G., Zhong,M. *et al.* (2016) DNA methylation on N(6)-adenine in mammalian embryonic stem cells. *Nature*, **532**, 329–333.
28. Huang,W., Xiong,J., Yang,Y., Liu,S.M., Yuan,B.F. and Feng,Y.Q. (2015) Determination of DNA adenine methylation in genomes of mammals and plants by liquid chromatography/mass spectrometry. *RSC Adv.*, **5**, 64046–64054.
29. Xiao,C.L., Zhu,S., He,M., Chen, Zhang,Q., Chen,Y., Yu,G., Liu,J., Xie,S.Q., Luo,F. *et al.* (2018) N(6)-Methyladenine DNA modification in the human genome. *Mol. Cell*, **71**, 306–318.
30. Schiffers,S., Ebert,C., Rahimoff,R., Kosmatchev,O., Steinbacher,J., Bohne,A.V., Spada,F., Michalakis,S., Nickelsen,J., Muller,M. *et al.* (2017) Quantitative LC–MS provides no evidence for m(6) dA or m(4) dC in the genome of mouse embryonic stem cells and tissues. *Angew. Chem. Int. Ed. Engl.*, **56**, 11268–11271.

31. Liang,Z., Shen,L., Cui,X., Bao,S., Geng,Y., Yu,G., Liang,F., Xie,S., Lu,T., Gu,X. *et al.* (2018) DNA N(6)-Adenine methylation in arabidopsis thaliana. *Dev Cell*, **45**, 406–416.
32. Luo,G.Z., Blanco,M.A., Greer,E.L., He,C. and Shi,Y. (2015) DNA N(6)-methyladenine: a new epigenetic mark in eukaryotes? *Nat. Rev. Mol. Cell Biol.*, **16**, 705–710.
33. Jia,G., Fu,Y., Zhao,X., Dai,Q., Zheng,G., Yang,Y., Yi,C., Lindahl,T., Pan,T., Yang,Y.G. *et al.* (2011) N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.*, **7**, 885–887.
34. Fu,Y., Jia,G., Pang,X., Wang,R.N., Wang,X., Li,C.J., Smemo,S., Dai,Q., Bailey,K.A., Nobrega,M.A. *et al.* (2013) FTO-mediated formation of N(6)-hydroxymethyladenosine and N(6)-formyladenosine in mammalian RNA. *Nat. Commun.*, **4**, 1798.
35. Xiao,C.L., Zhu,S., He,M.H., Chen,D., Zhang,Q., Chen,Y., Yu,G.L., Liu,J.B., Xie,S.Q., Luo,F. *et al.* (2018) N-6-Methyladenine DNA modification in the human genome. *Mol. Cell*, **71**, 306–318.
36. Muller,T.A., Struble,S.L., Meek,K. and Hausinger,R.P. (2018) Characterization of human AlkB homolog 1 produced in mammalian cells and demonstration of mitochondrial dysfunction in ALKBH1-deficient cells. *Biochem. Biophys. Res. Commun.*, **495**, 98–103.
37. Muller,T.A., Tobar,M.A., Perian,M.N. and Hausinger,R.P. (2017) biochemical characterization of AP lyase and m(6)A demethylase activities of human AlkB homologue 1 (ALKBH1). *Biochemistry*, **56**, 1899–1910.
38. Lu,K., Collins,L.B., Ru,H., Bermudez,E. and Swenberg,J.A. (2010) Distribution of DNA adducts caused by inhaled formaldehyde is consistent with induction of nasal carcinoma but not leukemia. *Toxicol. Sci.*, **116**, 441–451.
39. Luo,G.Z., Wang,F., Weng,X., Chen,K., Hao,Z., Yu,M., Deng,X., Liu,J. and He,C. (2016) Characterization of eukaryotic DNA N(6)-methyladenine by a highly sensitive restriction enzyme-assisted sequencing. *Nat. Commun.*, **7**, 11301.
40. Xing,X.W., Tang,F., Wu,J., Chu,J.M., Feng,Y.Q., Zhou,X. and Yuan,B.F. (2014) Sensitive detection of DNA methyltransferase activity based on exonuclease-mediated target recycling. *Anal. Chem.*, **86**, 11269–11274.
41. Liu,H. and Naismith,J.H. (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.*, **8**, 91.
42. Lu,K., Gul,H., Upton,P.B., Moeller,B.C. and Swenberg,J.A. (2012) Formation of hydroxymethyl DNA adducts in rats orally exposed to stable isotope labeled methanol. *Toxicol. Sci.*, **126**, 28–38.
43. Pfeifer,G.P., Xiong,W., Hahn,M.A. and Jin,S.G. (2014) The role of 5-hydroxymethylcytosine in human cancer. *Cell Tissue Res.*, **356**, 631–641.