8514–8531 *Nucleic Acids Research, 2023, Vol. 51, No. 16* *Published online 27 July 2023*
https://doi.org/10.1093/nar/gkad628

NAT10-mediated N4-acetylcytidine mRNA modification regulates self-renewal in human embryonic stem cells

Rucong Liu[1,2,†], Zibaguli Wubulikasimu[1,†], Runze Cai[1,†], Fanyi Meng[1], Qinghua Cui[2], Yuan Zhou[2,*] and Yang Li[1,*]

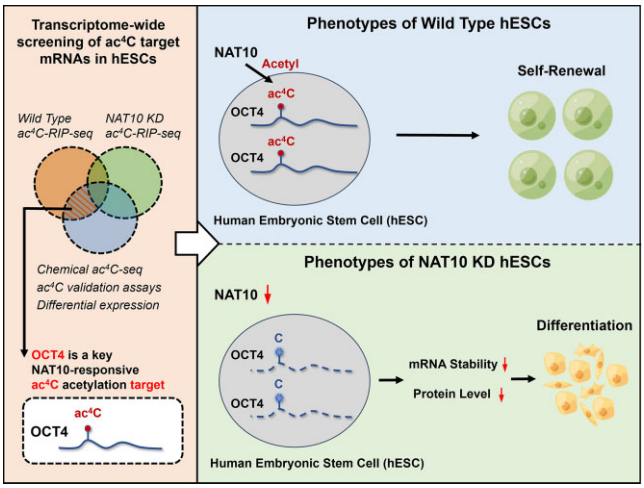
[1]Department of Cell Biology, School of Basic Medical Sciences, Peking University Stem Cell Research Center, Peking University, Beijing 100191, China and [2]Department of Biomedical Informatics, State Key Laboratory of Vascular Homeostasis and Remodeling, School of Basic Medical Sciences, Peking University, Beijing 100191, China

Received October 05, 2022; Revised July 05, 2023; Editorial Decision July 10, 2023; Accepted July 15, 2023

ABSTRACT

NAT10-catalyzed N^4-acetylcytidine (ac^4C) has emerged as a vital post-transcriptional modulator on the coding transcriptome by promoting mRNA stability. However, its role in mammalian development remains unclear. Here, we found that NAT10 expression positively correlates with pluripotency *in vivo* and *in vitro*. High throughput ac^4C-targeted RNA immunoprecipitation sequencing (ac^4C-RIP-seq), $NaCNBH_3$-based chemical ac^4C sequencing (ac^4C-seq) and liquid chromatography-tandem mass spectrometry (LC–MS/MS) assays revealed noticeable ac^4C modifications in transcriptome of hESCs, among which transcripts encoding core pluripotency transcription factors are favorable targets of ac^4C modification. Further validation assays demonstrate that genetic inactivation of NAT10, the ac^4C writer enzyme, led to ac^4C level decrease on target genes, promoted the core pluripotency regulator *OCT4 (POU5F1)* transcript decay, and finally impaired self-renewal and promoted early differentiation in hESCs. Together, our work presented here elucidates a previously unrecognized interconnectivity between the core pluripotent transcriptional network for the maintenance of human ESC self-renewal and NAT10-catalyzed ac^4C RNA epigenetic modification.

GRAPHICAL ABSTRACT



INTRODUCTION

Chemical modifications of messenger RNAs (mRNAs) are emerging as a key component of the post-transcriptional gene expression regulation machinery that involves multiple processes including degradation, splicing and translation of mRNAs. With specific writer and eraser enzymes responsible for the formation and removal of the modifications respectively, reversible mRNA modifications enable a direct and rapid way to remodulate cellular transcriptome for rapid adaptation to the changing environment (1–3). Accumulating evidence has indicated the extensive regulatory effects of RNA modifications on transcriptome fractions, especially those related to cell proliferation and differentiation. For example, N^6-methyladenosine (m^6A) has been identified as the most abundant internal modification of mRNAs in eukaryotes, linking external stimuli to an intricate network of transcriptional, post-transcriptional

*To whom correspondence should be addressed. Tel: +86 10 82802164; Fax: +86 10 82801001; Email: liyang@bjmu.edu.cn
Correspondence may also be addressed to Yuan Zhou. Tel: +86 10 82801585; Fax: +86 10 82801001; Email: zhouyuanbioinfo@hsc.pku.edu.cn
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

and translational processes (4). These m^6A modifications affect a broad spectrum of cellular functions, including maintenance of the pluripotency of embryonic stem cells (ESCs) and reprogramming of somatic cells into the induced pluripotent stem cells (iPSCs) (3,5,6).

Recently, N^4-acetylcytidine (ac^4C) has emerged as a vital post-transcriptional modulator on the coding transcriptome (7). Ac^4C is a highly conserved RNA modification in eukaryotic cells, it was first discovered in the yeast tRNA in 1966 (8). Recent studies have shown that the presence of ac^4C on tRNAs helps to increase the high fidelity of protein translation (9,10), and maintains the thermotolerance of the organism in Orita *et al.*'s study (11). The ac^4C modification is also found on rRNA. In 1978, Thomas *et al.* found ac^4C on eukaryotic 18S rRNA, indicating that ac^4C could be present in RNA types other than tRNA (12). In 2015, Sharma *et al.* found two ac^4C sites on the 18S rRNA of germinated fission yeast and human HCT116 cells, which is important for maintaining translation accuracy (13). Most of the early ac^4C studies focused on the observations from tRNAs and rRNAs, but very recently, a large number of ac^4C modifications have also been detected in human and yeast mRNAs (7,14). By transcriptome-wide ac^4C-targeted RNA immunoprecipitation sequencing (ac^4C-RIP-seq), Arango *et al.* found that ac^4C is widely distributed in mRNA coding sequences, promoting translation and mRNA stability (7). All documented ac^4C events in tRNA, rRNA and mRNA are catalyzed by the only known ac^4C writer, *N*-acetyltransferase 10 (NAT10) (in human) or its homolog Kre33 (in yeast) (13,15). Besides, an adaptor protein THUMPD1 can assist NAT10 and Kre33 in modifying tRNA with ac^4C acetylation (13,16). Furthermore, recent study has shown that small nucleolar RNA SNORD13 is specifically required for rRNA ac^4C acetylation (17).

In line with the wide scope of transcripts targeted by NAT10-catalyzed ac^4C modification, ac^4C has been demonstrated to be involved in the occurrence of various diseases including cancer, inflammatory, heart and metabolic diseases (18,19). However, its functions in mammalian development have not been clarified. Human ESCs (hESCs) have the potential of indefinite self-renewal and the ability to differentiate into any cell type of all three germ layers, which makes them become an excellent *in vitro* system to study mammalian development and disease modeling (20). To address potential roles for ac^4C in self-renewal and differentiation, we generated NAT10-, THUMPD1- and SNORD13-deficient hESCs by using the CRISPR/Cas9 gene-editing (Knockout, KO) and lentivirus-mediated RNAi (Knockdown, KD) approaches. High throughput ac^4C-RIP-seq, chemical ac^4C-seq and LC–MS/MS assays showed that in hESCs, ac^4C modifications were substantially presented on mRNAs. Phenotypic assays showed that deficiency of ac^4C writer NAT10, but not its key co-factors for tRNA and rRNA ac^4C modifications (i.e. THUMPD1 and SNORD13), significantly perturb hESC self-renewal and proliferation. Together, our findings demonstrated that NAT10-mediated ac^4C mRNA modification is required for maintenance of self-renewal of hESCs.

MATERIALS AND METHODS

Cell lines and culture conditions

Human ESC lines H1 (WiCell Research Institute) were cultivated in serum-free, defined TeSR-E8 medium (STEMCELL Technologies, Vancouver, Canada) and grown on plates coated with growth-factor-reduced Matrigel (Corning, NY, USA). hESCs were passaged with non-enzymatic dissociation by 0.5 mM EDTA solution (Cellapy, Beijing, China).

Real-time qPCR for mRNA quantification

Total RNA was prepared with TRIzol (Invitrogen, Carlsbad, CA, USA), and 500 g RNA was transcribed into cDNA using Primescript RT master mix (Takara, Japan). Real-time qPCRs were performed on Bio-Rad CFX384 real-time PCR detection system using PowerUp SYBR Green Master Mix (ThermoFisher Scientific, Boston, MA, USA). GAPDH was used as internal control to normalize the data across different samples. All primer sequences for real time qPCR are listed in Supplementary Table S1.

Western blot

Cells were lysed on ice using RIPA lysis buffer (Beyotime, Shanghai, China) and centrifuge for 15 min at 12 000 × g at 4°C. Protein amounts were quantified using the BCA Protein Assay Kit (Beyotime). Proteins were separated by 10% SDS-PAGE, transferred onto PVDF membranes (Merck-Millipore, Germany). Membranes were blocked with 5% skimmed milk (Beyotime) for 1 h and probed overnight with primary antibodies (summarized in Supplementary Table S2) at 4°C overnight, followed by incubation with HRP-conjugated secondary anti-IgG antibody (Beyotime) for 1 h at room temperature. Membranes were detected with HRP substrate luminol reagents (Merck-Millipore), and scanned using the Gel Doc EZ Imager (Bio-Rad). Quantification of blot intensity was performed using ImageJ software (NIH, MD).

CRISPR-mediated NAT10, THUMPD1 and SNORD13 KO

The CRISPR-Cas9 knock-out vector, containing cassettes for GFP, Cas9 endonuclease, and the double guide RNA (gRNA) moiety designed to target NAT10, THUMPD1 (21) or SNORD13 (17), was custom synthesized (GenePharma, Suzhou, China). After transfection with Lipofectamine 3000 (Life Technologies), GFP^+ hESCs were sorted using flow cytometry, and plated at low density for single-colony isolation. Selected single colonies were tested by western blot for loss of protein and RT-qPCR for loss of RNA. Mutations were validated by sequencing products of PCR amplification of the regions flanking the targeting sites. The CRISPR target sequences and primer sequences for RT-qPCR analysis of SNORD13 RNA are included in Supplementary Tables S3-S4.

RNAi-mediated NAT10 KD

Cells transduced with NAT10 shRNA or negative control shRNA lentivirus with GFP (Genechem, Shanghai, China) were cultured for 48 h. GFP+ cells then were enriched and expanded in TeSR-E8 medium with 0.5 μg/ml puromycin. The target sequences of shRNA are included in Supplementary Table S3.

Lentivirus-mediated NAT10 overexpression

Lentiviral particles overexpressing NAT10 (NAT10 OE) and empty vector (control) were purchased from Genechem corporation (Shanghai, China). For generation of NAT10 OE hESCs, hESCs were transduced with lentiviral particles expressing the complete nucleotide sequence of NAT10. 48 h after transduction, NAT10 OE and control cells were enriched and expanded in TeSR-E8 medium with 0.5 μg/ml puromycin.

Cell counting kit 8 (CCK8) assay and ALP staining

Cells were digested into single cell suspension using Accutase (Merck-Millipore) and counted in a counting plate. After counting, the cells were seeded in Matrigel-coated 96-well with a cell density of 1000 cells per well. After appropriate cultivate time, 10% CCK8 (Dojindo, Kyushu, Japan) was added into the culture medium and incubate cells for 2 h, then optical density value at 450 nm was detected using a microplate reader. Cells were detected in every 24 h.

For ALP (alkaline phosphatase) staining, the cells were plated on Matirgel-coated 12-well at same count and incubate for 3 days. The cells were fixed with 4% paraformaldehyde for 3–5 min followed by washing with PBS. Then, the samples were added with BCIP/NBT staining solution (Beyotime) for 15 min in dark at room temperature. The usage of regents was followed with the manufacturer's recommendations. The cell images were taken after washing the cells with PBS.

Spontaneous differentiation via EB formation

To generate EBs, hESCs were dissociated into small clumps, then plated onto Ultra low attachment surface 6-well plate (Corning). The suspended cell clumps were cultured in DMEM/F12 supplemented with L-glutamine (Gibco, Carlsbad, CA, USA) and 15% fetal bovine serum (Gibco) for 6 days.

Definitive endodermal cell differentiation

To initiate definitive endodermal cell differentiation, hESCs were cultured on Matrigel-coated-12-well plate for 3 days in RPMI1640 (Gibco) /B27 supplement (Gibco) and 100 ng/ml Activin A (Peprotech, NJ, USA). After 3 days, endoderm markers were detected by flow cytometry analysis.

Neural ectodermal cell differentiation

Monolayer culture protocol was used to initiate neural differentiation. hESCs were plated onto Matrigel-coated 6-well plates at a cell confluence of 2 × 105 cells/cm^2, and then cultured in Neural induction medium (Stem Cell Technologies), changing to fresh culture medium every day. The cells were collected for flow cytometry analysis after 3 days.

Teratomas formation

Animal experiments were performed in accordance with guidelines of the Institutional Animal Care and Use Committee (IACUC) at Peking University. Male NOD-SCID mice (4–6 weeks of age) were purchased from the Department of Experimental Animal, Peking University Health Science Center. hESCs were harvested by EDTA solution treatment, suspended in PBS, and subcutaneously injected into the thighs of mice. After about 6 weeks, teratomas were explanted, fixed in 4% paraformaldehyde, embedded in paraffin, and examined histologically using immunohistochemistry staining.

Total RNA isolation and poly(A)-RNA purification

Total RNA was extracted using TRIzol (Invitrogen). mRNA was isolated by two rounds of purification using the Dynabeads® mRNA Purification Kit (Thermo Fisher Scientific, MA, United States) according to the manufacturer's protocol. The purity of the isolated poly(A)-RNA (mRNA fraction) was verified by RT-qPCR using specific primers for the detection of 18S rRNA and 28S rRNA and leucine/serine tRNA that are known to have ac^{4}C modifications (13). The primer sequences for RT-qPCR analysis of 18S rRNA and 28S rRNA are listed in Supplementary Table S4. The commercial primer sequences for leucine/serine tRNAs are proprietary to Ribo Life Science Co., Ltd (Guangzhou, China).

Acetylated RNA immunoprecipitation sequencing (ac^{4}C-RIP-seq) and data analysis

The ac^{4}C-RIP-seq was performed by CloudSeq Biotech Co., LTD (Shanghai, China). Anti-ac^{4}C immunoprecipitation reaction (ac^{4}C-RIP) was performed using GenSeq's RNA acetylation kit. RNA was randomly segmented into about 200 nt segments. Protein A/G magnetic beads (Thermo Fisher Scientific) and ac^{4}C antibody (Abcam, Cambridge, United Kingdom) or control rabbit IgG antibodies (GenSeq, Shanghai, China) were rotated and incubated at room temperature for 1 hour to bind the antibody to the magnetic beads. Then the RNA fragments and the antibody bound with magnetic beads were rotated and incubated at 4°C for 4 hours. RNA/antibody complexes were eluted from magnetic beads by Protease K and RNA was extracted by phenol-chloroform method. RNA sequencing libraries were constructed using the NEBNext® Ultra II Directional RNA Library Prep Kit (Thermo Fisher Scientific). The constructed libraries were quality-controlled with Agilent 2100 bioanalyzer and then sequenced on Illumina NovaSeq 6000 sequencer (Illumina, Inc., San Diego, CA, USA) for high-throughput sequencing. After sequencing, image analysis, base identification and quality control, raw reads (FASTQ raw data) were obtained.

Regarding to the ac^{4}C-RIP-seq data processing, fastp software (v0.21.1) (22) was first used to remove the adaptor and low-quality reads to obtain high-quality clean

FASTQ reads. Clean reads were then aligned to the reference genome (hg38) by using HISAT2 software (v2.2.1) (23), with the genome annotation from Ensembl (http://www.ensembl.org, version 107) as the reference gene model. The aligned SAM-format reads were sorted, transformed to BAM format, and indexed via samtools (v1.12) (24). Finally, the BAM-format reads for each paired IP + input (or IgG + input) samples were processed by exomePeak2 software (http://bioconductor.org/packages/release/bioc/html/exomePeak2.html, v1.8.1) to identify significant acetylation peaks in each IP or IgG sample. Compared to canonical peak calling methods like MACS2, exomePeak and its updated version exomePeak2 are dedicated methods for RNA modification peak calling that take known exon topologies into consideration, enabling a more accurate peak identification from splicing-aware RNA read data (25). We also noted a better between-sample repetition rate with exomePeak2 than that with MACS2 for ac^4C acetylation peak calling. The exomePeak2 software was implemented with default parameter and the same Ensembl v107 reference gene model. To control false positive, any IP peak showing more than 10% coordinate overlap with the non-specific IgG peaks were removed by the *subtract* function of bedtools (v2.30.0) (26). Moreover, peaks that cannot be mapped to known genes, and peaks on the less confident acetylation target genes (total peak score < 20) were discarded.

To better annotate the acetylation peaks, HOMER software (27) was used for de novo discovery of the acetylation motif. Distribution of ac^4C peaks along the transcripts was characterized by R package Guitar (28). Interval annotation analysis of ac^4C peaks along the transcripts was characterized by R package ChIPseeker (29). Visualization of the ac^4C peaks as per-base coverage across the genome was performed via the Integrative Genome Viewer (IGV) (30). Gene Ontology (GO) functional enrichment analysis was performed by R (v4.2.2) and GO terms with *q*-values [i.e. *P*-values corrected by the *q*-value method (31)] smaller than 0.05 were considered statistically significant terms.

$NaCNBH_3$-based chemical ac^4C sequencing (ac^4C-seq) and data analysis

High-throughput $NaCNBH_3$-based chemical ac^4C sequencing (ac^4C-seq) was conducted by Seqhealth Technology Co., Ltd (Wuhan, China). Total RNA was extracted from wild-type (WT) and NAT10 KD hESCs using TRIzol Reagent (Invitrogen). DNA digestion was carried out after RNA extraction by DNase I. RNA quality was determined by examining A260/A280 with Nanodrop spectrophotometer (Thermo Fisher Scientific). RNA Integrity was confirmed by Qsep100 (bioptic) or Agilent5300 (Agilent). 30 μg total RNA was used for mRNA capture using KAPA mRNA Capture Kit (Beijing, China) following the manufacturer's instruction. A fraction of captured RNA was left as the input samples, and the other was treated with $NaCNBH_3$ at 25°C for 20 min; RNA clean beads were used to retrieve the treated RNA. Then, KC Digital Stranded mRNA Library Prep Kit for Illumina (Wuhan, China) was used for library construction. The kit eliminates duplication bias in PCR and sequencing steps, by using unique molecular identifier (UMI) of 8 random bases to label the pre-amplified cDNA molecules. The library products corresponding to 200–500 bp were enriched, quantified and finally sequenced on Illumina Novaseq 6000 with PE150 model.

After sequencing, both total low-quality reads and adaptor sequences were first trimmed by fastp (v0.21.1) (22). The in-house software of Seqhealth corporation (Wuhan, China) was used to eliminate duplicated reads according to UMI tags, and the UMI tag sequences were also trimmed by process. Introduction of UMI tags and read de-duplication by UMI tags are helpful to remove PCR duplication artifacts, permitting more accurate mutation and variant calling. Clean reads after duplication removal were then aligned to reference genome (hg38) by using STAR software (v2.7.8a) (32), with the same Ensembl v107 genome annotation as the reference gene model. The aligned SAM-format reads were sorted, transformed to BAM format, and indexed via samtools (v1.12) (24). The JACUSA software (version 2.0.0-S-BETA-1) (33) was used for detecting single nucleotide variant (SNV) sites and obtaining the corresponding pileups. Finally, the standard algorithm code for ac^4C-seq modification site identification (https://github.com/SchwartzLab/ac4c-seq) (34) was utilized to identify the modification sites based on the significantly overrepresented SNV events at the ac^4C sites.

RNA sequencing (RNA-seq) and data analysis

The transcriptome profiling of NAT10 KD and sh control hESCs were performed via RNA-seq. Total RNA was isolated and reverse transcribed into cDNA to generate indexed Illumina library, followed by sequencing at the Beijing Genomics Institute (Beijing, China) using the BGISEQ-500 platform. The raw reads were quality controlled, processed, and alignment quality-controlled, and aligned to reference genome following the same protocol as what was used in ac^4C-RIP-seq data processing. FeatureCounts (v2.0.3) (35) software was used for gene expression quantification. Differential gene detection was performed using DESeq2 (v1.26.0) (36). An adjusted *P*-value < 0.05 and a fold change (FC) > 1.5 or $< 2/3$, as per the DESeq2 calculation results, were used as the threshold to identify differentially expressed genes (DEGs). The principal component analysis (PCA) and GO functional enrichment analysis were performed by R (v4.2.2). GO terms with *q*-value < 0.05 were considered statistically significant terms.

Transcriptome-wide mRNA half-life profiling and data analysis

Actinomycin D (Sigma-Aldrich) was added to the medium with 2 μg/ml. Cells were harvested at 0, 1.5 and 3.5 h after the addition of actinomycin D. Total RNA was isolated and mRNA was purified from total RNA using oligo (dT)-attached magnetic beads. The resulting mRNA molecules were then fragmented and reverse-transcribed into double-stranded cDNA. ERCC spike-in (Invitrogen) was added to total RNA before library construction, following the protocol recommended by ERCC spike-in kit manufacturer. RNA sequencing was performed by Beijing Genomics Institute (Beijing, China) using BGISEQ-500 platform.

After sequencing, raw data was quality controlled, processed, and alignment quality-controlled, and aligned to reference genome following the same protocol as what was used in ac^{4}C-RIP-seq data processing. FeatureCounts (v2.0.3) (35) software was used for gene expression quantification. After quantification, absolute transcript concentrations per cell were computed by linear-fitting the RNA spike-in TPM values to attomoles, and RNA half-life measures were calculated based on this attomole expression quantification, as described by Wang *et al.* (2). In short, for every transcript, the RNA-degradation constant was determined as the $\log_2$ fold change in RNA abundance 1.5 or 3.5 h after transcription was inhibited, and the average degradation constant was used to calculate the RNA half-life per transcript. The heatmap illustration of half-life changes was drawn using the R package ComplexHeatmap (v2.14.0) (37).

Liquid chromatography–tandem mass spectrometry (LC–MS/MS)

Both total RNA and mRNA samples were analyzed by LC–MS/MS to assess the corresponding ac^{4}C quantification. For mRNA samples, rRNA/tRNA contaminations were tested by RT-qPCR before LC–MS/MS, as described above. RNA modification assay by LC–MS/MS was performed by CloudSeq Biotech Co., LTD (Shanghai, China). Briefly, 2 μg of total RNA or mRNA was digested into nucleosides with 2 U of nuclease P1 (Sigma) and 1.5 U of alkaline phosphatase (Sigma-Aldrich) at 37°C for 3 h. Then, 5 μl of the mixture solution was injected into an LC–MS/MS system. The nucleosides were detected by an AB Sciex QTRAP 6500 LC–MS/MS platform in positive ion mode. AB 6500 QTRAP LC–MS/MS System, equipped with an ESI Turbo Ion-Spray interface, operating in both positive and negative ion modes and controlled by Analyst 1.6.3 software (AB Sciex). The ESI source operation parameters were as follows: ion source, turbo spray; source temperature 550°C; ion spray voltage (IS) 5500 V. DP and CE for individual MRM transitions was done with further DP and CE optimization. A specific set of MRM transitions were monitored for each period according to the RNA eluted within this period.

Dot blot assay of RNA ac^{4}C acetylation

The dot blot assay was performed by Genelily Biotechnology Co., LTD (Shanghai, China). Briefly, total RNA was denatured at 65°C for 5 min to disrupt the secondary structure, followed by immediately cooling on ice to prevent re-formation of the secondary structure. 2 μl of RNA was dropped directly onto Immobilon®-Ny + Membrane (Merck-Millipore) and cross-linked in UV crosslinkers. Unbound mRNA was washed with gentle shaking for 5 min. Next, membranes were incubated in 10 ml blocking buffer (5% skim milk powder) for 1 h at room temperature with gentle shaking and then placed in 5 ml anti-ac^{4}C antibody dilution buffer (1:250 dilution; 2 μg/ml) at 4°C overnight. Membranes were next washed 3 times in wash buffer for 10 min each time and then placed in 5 ml Goat anti rabbit IgG-HRP (1:1000 diluted; 20 ng/ml) (Beyotime) secondary antibody dilution buffer for 1 h at room temperature. Membranes were gently washed 3 times in wash buffer for 10 min each time and then incubated in 3 ml of western blotting substrate for 5 min at room temperature in the dark. Dot blot results were recorded by Tanon automatic chemiluminescence image analysis system (Beijing, China). For internal standard detection, membranes were incubated with 0.02% methylene blue (Sigma-Aldrich) in 0.3 M sodium acetate (pH 5.2) for 10–15 min and then washed with ddH_2O.

Ac^{4}C-RIP-qPCR

ac^{4}C-RIP was conducted by an RNA Immunoprecipitation Kit (GENESEED, Guangzhou, China). In brief, cells were washed with PBS twice and collected by centrifugation at 1000 × g for 5 min. Then 1 ml of RIP lysis buffer was added to the cells on ice for 10 min. Then 100 ml of lysis buffer was stored at –80°C. An anti-ac^{4}C antibody or normal rabbit IgG (Cell Signaling Technology, MA, United States) was mixed with protein A/G beads at 4°C for 2 h. Then 450 μl of lysis buffer was incubated with the beads above at 4°C for 2 h. The beads were washed with buffer, and RNA was extracted. The RNAs were examined using RT-qPCR with specific primers and the results of qPCR data were analyzed according to the manufacturer's instruction. That is, the relative expression of target genes was calculated based on the $2^{-\Delta\Delta Ct}$ method, where the $\Delta\Delta Ct$ term is denoted as:

$$\Delta\Delta Ct = \left(Ct_{IP} - Ct_{Input}\right) - \left(Ct_{IgG} - Ct_{Input}\right)$$

where the Ct_{IP}, Ct_{Input} and Ct_{IgG} values are the cycle thresholds (Ct) of the IP group, Input group, and IgG group, respectively. Of note, the Ct_{Input} value of the target genes has been cancelled out, and the relative expression in ac^{4}C-RIP-qPCR is only determined by Ct_{IP} and Ct_{IgG}. For IgG group, since Ct_{IP} is the same as Ct_{IgG}, the relative expression in IgG group is always $2^0 = 1$. As for IP group, the final result is the relative fold enrichment in IP group comparing to IgG group. And all primer sequences for ac^{4}C-RIP-qPCR are listed in Supplementary Table S5.

RNA stability assay by RT-qPCR

Actinomycin D (Sigma-Aldrich) at 2 μg/ml was added to hESCs 24 h after they were seeded in a 6-well plate. After 0, 3 or 6 h of incubation, cells were collected and RNAs were isolated for RT-qPCR.

Statistical analysis

Unless otherwise stated, bar charts represent the mean ± standard deviation of the mean (SD), and the results for the two groups were compared using a two-tailed unpaired Student's *t*-test. A *P*-value smaller than 0.05 was considered significant unless stated differently, and the exact degree of significance as indicated by asterisks is stated in the legends. Statistical significance was presented as $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$; ns indicates $P > 0.05$.

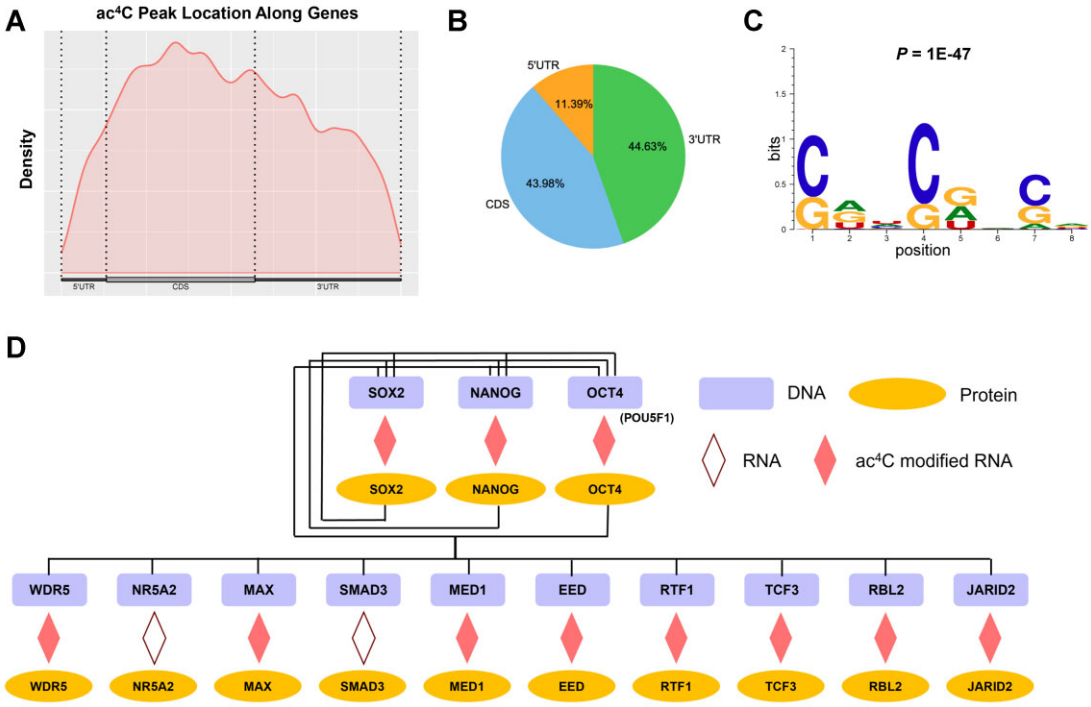


Figure 1. Overview of ac^{4}C Modification Pattern by ac^{4}C-RIP-Seq. (**A**) Normalized distribution of ac^{4}C peaks along 5′UTR, CDS and 3′UTR of mRNAs in WT hESCs. (**B**) Pie chart showing fractions of ac^{4}C peaks in different mRNA regions in WT hESCs. (**C**) Sequence logo of enriched motifs within ac^{4}C peaks in WT hESCs. (**D**) Diagram showing the ac^{4}C modification for genes involved in core pluripotency transcriptional network (the network topology is adapted from Young, 2011; permission of reuse has been acquired from Elsevier, with the License Number of 5587090446553).

RESULTS

Prevalent ac^{4}C modification in hESC transcriptome revealed by high-throughput sequencing

To explore the transcriptome-wide profile of ac^{4}C modification, we mapped the locations of ac^{4}C modification on wild-type (WT) hESCs by high-throughput ac^{4}C RNA immunoprecipitation sequencing (ac^{4}C-RIP-seq). As the result, 37 988 ac^{4}C peaks covering 6916 genes and 38 150 ac^{4}C peaks covering 7350 genes were identified in two hESC replicates, respectively. We observed that the ac^{4}C peaks were also enriched in coding sequences (CDS), but relatively less enriched in 5′-UTRs than previous observations in HeLa cells (7) and a considerable fraction of ac^{4}C acetylation peaks also settled in 3′-UTRs (Figure 1A-B). Furthermore, *de novo* motif analysis of hESCs ac^{4}C sites specifically identified a CNNCNNCN ac^{4}C sequence motif (Figure 1C), which is consistent with the motif previously observed in HeLa cells (7). Of note, among those mRNAs with ac^{4}C modification, we observed mRNAs encoding the core pluripotency regulators (38,39) *OCT4 (POU5F1)*, *SOX2* and *NANOG* were consistently modified with ac^{4}C. Besides, other pluripotency-related gene transcripts were also detected to have ac^{4}C modification, including *RTF1*, *RBL2*, *JARID2*, *WDR5*, *MED1*, *EED*, *MAX* and *TCF3* (Figure 1D). The 11 ac^{4}C-modifed core pluripotency regulators listed in Figure 1D were further compared with the previous ac^{4}C target genes identified by Arango *et al.* in HeLa cell. Five genes (*RTF1*, *JARID2*, *WDR5*, *EED* and *TCF3*) were also observed to have ac^{4}C modification in HeLa cells, while six genes (*OCT4 (POU5F1)*, *SOX2*, *RBL2*, *MED1*, *MAX* and *NANOG*) were not observed in HeLa cell but found ac^{4}C modification in hESCs, suggesting at least partial cell-type specificity of RNA ac^{4}C modification.

Acetyltransferase NAT10 expression positively correlates with pluripotency

In order to understand the role of ac^{4}C modification in pluripotency maintenance, we focused on NAT10, the only known ac^{4}C acetyltransferase in humans whose expression levels directly affect the formation of ac^{4}C on RNAs. We first investigated NAT10 expression pattern during human early embryonic development by referring to a comprehensive public gene expression dataset in human early embryogenesis (40). As shown in Figure 2A, high expression of NAT10 is observed in epiblasts, but its expression is prominently downregulated during three-germ layer development. We then carried out spontaneous differentiation assay by EB (embryoid body) formation to investigate NAT10 expression pattern in hESC differentiation. Indeed, we observed a marked decrease of NAT10 expression during spontaneous differentiation in mRNA and protein levels, which has similar tendency with OCT4 expression (Figure 2B–D). Moreover, the assays of directed differentiation towards neural-ectodermal and endodermal cells revealed a down-regulation of NAT10 and OCT4 during early differentiation, along with up-regulation of differentiation markers SOX1 and SOX17 in both mRNA and protein levels (Figure 2E–J).

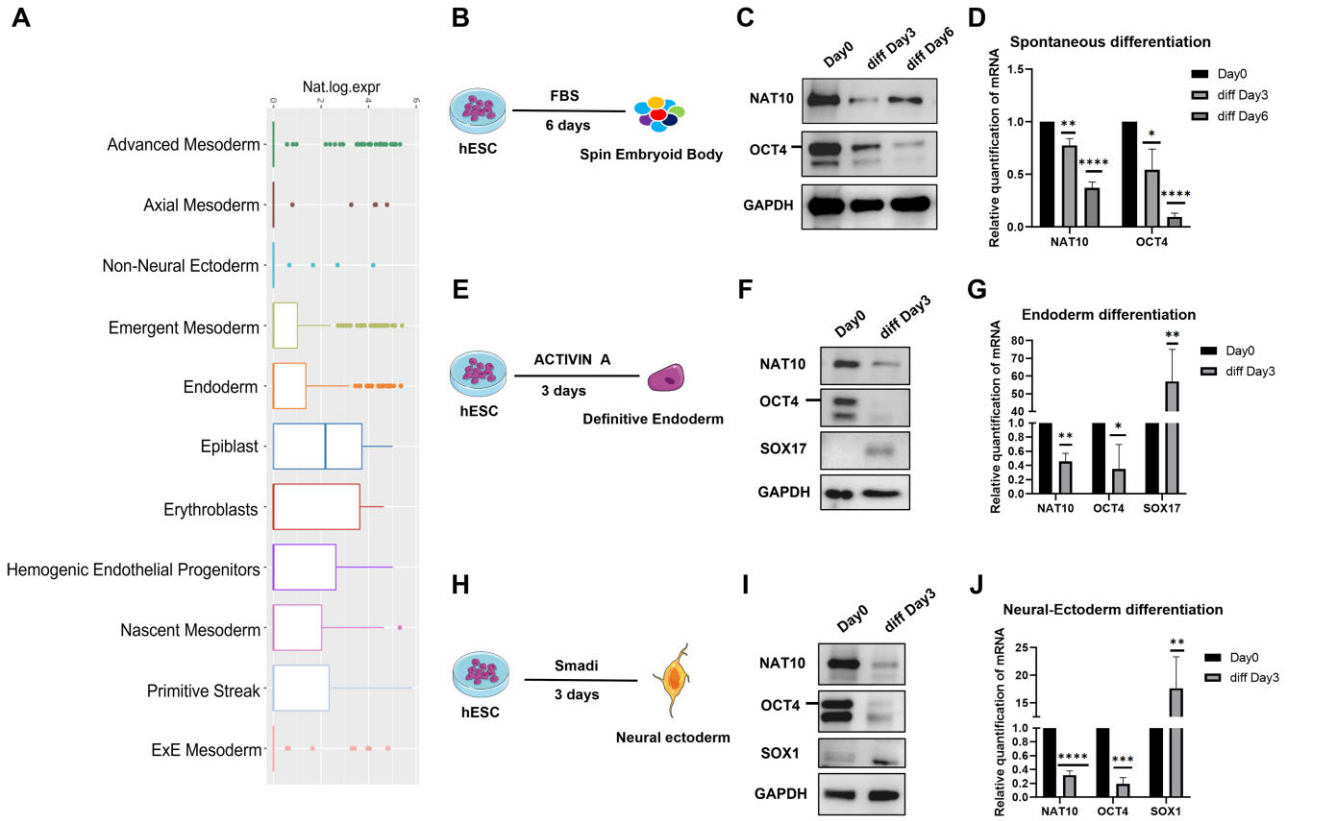


Figure 2. Correlation Between ac^{4}C Writer NAT10 and Pluripotency. (**A**) NAT10 expression pattern during human early embryonic development. (**B**) Experimental overview of embryoid body (EB) formation assay. (**C**, **D**) Representative western blot and RT-qPCR analysis for NAT10 and OCT4 expression on day 0, day 3 and day 6 of spontaneous differentiation by EB formation. OCT4 is the upper band marked by the black line. * means $P < 0.05$, ** means $P < 0.01$, **** means $P < 0.0001$. (**E**) Experimental overview of endodermal cell differentiation assay. (**F, G**) Representative western blot and RT-qPCR analysis for NAT10, OCT4 and SOX17 expression on day 0 and day 3 of endodermal cell differentiation. OCT4 is the upper band marked by the black line. * means $P < 0.05$, ** means $P < 0.01$. (**H**) Experimental overview of neural ectodermal cell differentiation assay. (**I, J**) Representative western blot and RT-qPCR analysis for NAT10, OCT4 and SOX1 expression on day 0, day 3 of neural ectodermal cell differentiation. OCT4 is the upper band marked by the black line. ** means $P < 0.01$, *** means $P < 0.001$, **** means $P < 0.0001$. Images in (C), (F) and (I) are representative of three independent experiments. In (D), (G) and (J), error bars represent mean ± SD ($n = 3$ independent experiments).

Generation of NAT10-deficient hESCs

To gain further insights into the role of NAT10 in hESCs, we generated NAT10 depletion hESCs. We first sought to knock out (KO) NAT10 protein expression by targeting the *NAT10* locus using the CRISPR/Cas9 gene-editing approach (Supplementary Figure S1A, B). After transfection with NAT10 KO vectors with GFP, single GFP$^+$ cells were sorted using flow cytometry, and plated into individual wells of 12-well plate for expansion. We observed two clones (clone 4# and 9#) could not form ES-like colony after sorting, showing complete differentiation morphology (Supplementary Figure S1C). It took very long period for us to harvest these cells for western blot to test NAT10 expression. We confirmed NAT10 expression was absent from two clones (Supplementary Figure S1D). After sorting, NAT10 KO cells could grow to 80–90% confluence in a well of 12-well plate, and then stopped proliferation. Except for western blot assay, no residual KO cells were available for further assays. NAT10 KO experiments were repeated for several times, as a result, we were unable to obtain stable, undifferentiated NAT10 KO hESCs.

Considering that the complete deletion of NAT10 expression affects cell survival, our study switched to RNAi to generate NAT10 knockdown (KD) cells. Western blot and RT-qPCR analysis confirmed a significant decrease of NAT10 protein and mRNA in hESCs transduced with NAT10 shRNA compared to control shRNA lentivirus (Supplementary Figure S2A–C). Dot blot analysis showed that there was a significant decrease of ac^{4}C levels in total RNA after NAT10 deficiency (Supplementary Figure S2D).

Perturbation of NAT10 impairs hESC self-renewal and promotes early differentiation in hESCs

We observed that NAT10 KD hESCs and sh control hESCs exhibited distinct morphology and size. The NAT10 KD hESCs exhibited a gradual and spontaneous differentiation morphology losing smooth edge. What's more, KD cells exhibited less and smaller colonies than sh control cells after planting at the same density and cultured under the same condition (Figure 3A). Consistent to the colony forming results, the CCK8 and ALP staining assays revealed that NAT10 KD hESCs displayed a significant lower

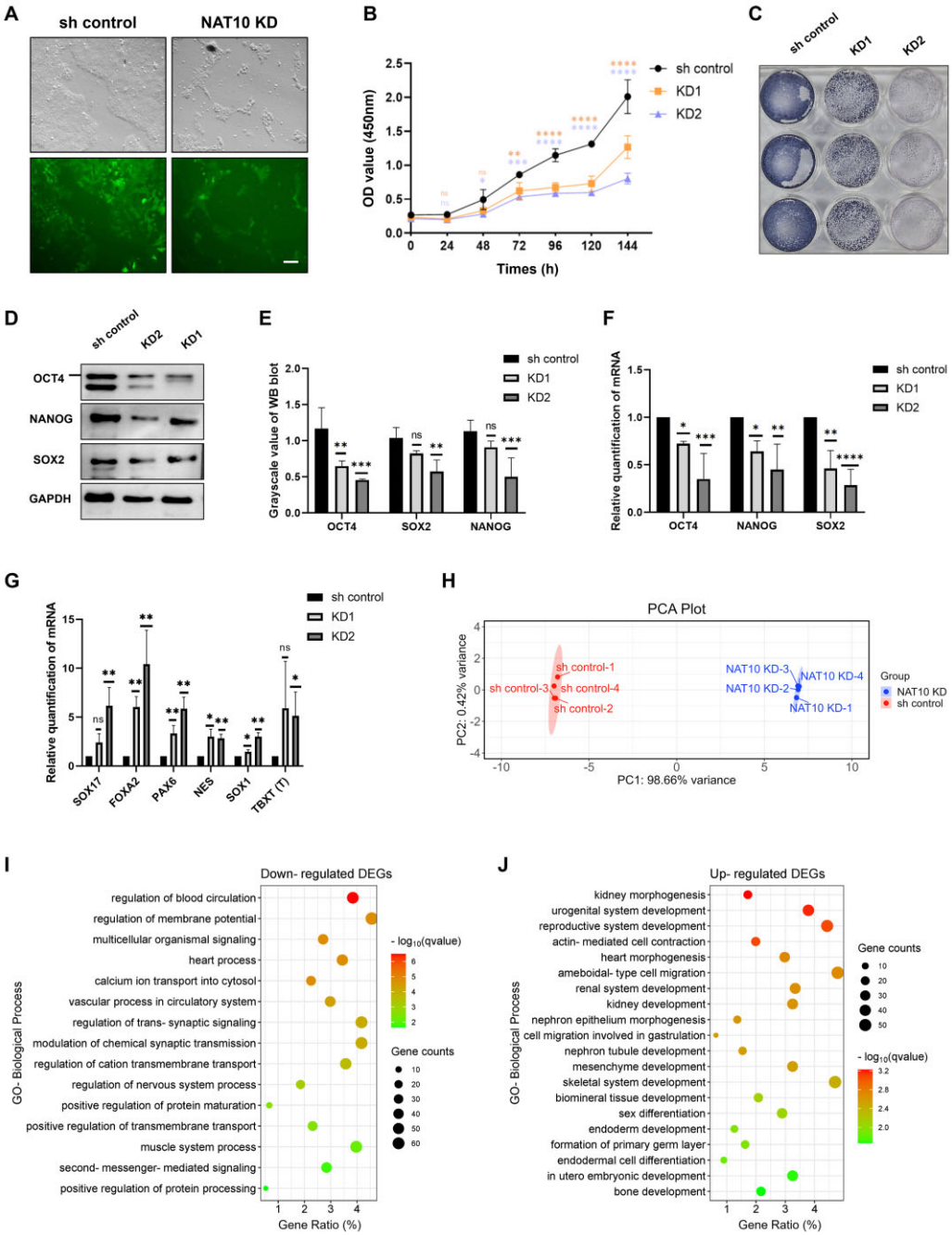


Figure 3. Impact of NAT10 KD on proliferation and differentiation of hESCs. (**A**) Representative images of morphology of sh control versus NAT10 KD cells under fluorescent and bright field. Scale bar: 50 μm. (**B**) Growth curve assessing cell proliferation kinetics of NAT10 KD versus sh control cells. * means $P < 0.05$, ** means $P < 0.01$, *** means $P < 0.001$, **** means $P < 0.0001$. ns means not significant. (**C**) Alkaline phosphatase (ALP) staining for sh control and NAT10 KD cells. (**D**) Representative western blot analysis of expression of OCT4, NANOG and SOX2 in sh control and NAT10 KD cells. OCT4 is the upper band marked by the black line. (**E**) The statistical histogram of grayscale value of western blot. ** means $P < 0.01$, *** means $P < 0.001$. ns means not significant. (**F**) RT-qPCR analysis of OCT4, SOX2, NANOG mRNA expression in sh control and NAT10 KD cells. * means $P < 0.05$, ** means $P < 0.01$, *** means $P < 0.001$, **** means $P < 0.0001$. (**G**) RT-qPCR analysis of definitive endoderm markers SOX17, FOXA2; neural-ectoderm markers SOX1, PAX6, NES (also known as NESTIN); and mesoderm marker TBXT (also known as T) in sh control and NAT10 KD cells. * means $P < 0.05$, ** means $P < 0.01$, ns means not significant. (**H**) Principal component analysis (PCA) plot depicting the clustering of NAT10 KD group versus sh control group. Each symbol represents an RNA-seq sample, and NAT10 KD and sh control samples are shown in blue and red, respectively. Sample groups with similar gene expression profiles are clustered with the indicated colors. The proportions of variation explained by first (PC1) and second (PC2) principal component are 98.66% and 0.42%, respectively. (**I**) Enriched functional terms for Gene Ontology Biological Processes (GO-BP) of the down-regulated DEGs in NAT10 KD hESCs. GO terms with q-value (i.e. *P*-value corrected by q-value method) smaller than 0.05 were considered statistically significant. The significantly enriched terms that are related to signal regulation pathways are shown here. (**J**) Enriched functional terms for GO-BP of the up-regulated DEGs in NAT10 KD hESCs. GO terms with q-value < 0.05 were considered statistically significant. The significantly enriched terms that are related to cell development and cell differentiation are shown. Images in (A), (C) and (D) are representative of three independent experiments. In (B), (E), (F) and (G), error bars represent mean ± SD ($n = 3$ independent experiments).

proliferation rate compare to sh control hESCs (Figure 3B, C). Next, we employed western blot and RT-qPCR analysis to detect gene expression of pluripotent marker and developmental regulators. NAT10 KD hESCs showed decrease in expression of core pluripotency regulators *OCT4, NANOG, SOX2* compared to sh control hESCs (Figure 3D-F), whereas increase in developmental regulators of three-germ layers, including endoderm markers *SOX17*, *FOXA2*, neural-ectoderm markers *SOX1*, *PAX6*, *NES (also known as NESTIN)* and mesoderm marker *TBXT (also known as T)* (Figure 3G), indicating that NAT10-deficiency induced the advanced differentiation in NAT10 hESCs.

Then we treated the NAT10 KD and sh control hESCs with STEMdiff Neural Induction medium and ACTIVIN A for directed differentiation into early neural ectoderm and definitive endoderm separately. After neural induction for 3 days, more $SOX1^+$ cells were present in the NAT10 KD cultures compared to sh control group (Supplementary Figure S3A-B). Consistent with tendency of neural differentiation, $FOXA2^+$ $SOX17^+$ double positive cells were significantly increased in the NAT10 KD cultures after endodermal induction for 3 days (Supplementary Figure S3C, D). The assays for directed differentiation demonstrated that NAT10 KD promoted early neural and endodermal development. Teratomas formation is an important indicator for evaluating *in vivo* differentiation ability of ESCs. To explore the effect of NAT10 KD on *in vivo* differentiation, NAT10 KD and sh control hESCs were injected subcutaneously into immunocompromised (NOD-SCID) mice to allow for teratomas development, and teratomas formation was observed 6 weeks later. Teratomas formed from NAT10 KD hESCs were significantly smaller in size and weight compare to sh control hESCs (Supplementary Figure S3E, F), which further confirmed the lower proliferation rate of NAT10 KD hESCs. The immunohistochemistry histological analysis revealed teratomas from both NAT10 KD and sh control hESCs containing all three germ layers (Supplementary Figure S3G).

In order to more comprehensively describe the alterations in hESC gene expression resulting from NAT10 knock-down, we performed high throughput RNA sequencing in NAT10 KD and sh control hESCs with 4 replicates for each group. The transcriptome profiling results showed that the knock-down of NAT10 resulted in down-regulation of 2455 genes and up-regulation of 1548 genes, respectively. PCA result showed that there were obvious distinctions between the expression profiles of NAT10 KD hESCs and those of sh control hESCs (Figure 3H). Gene Ontology Biological Processes (GO-BP) functional enrichment analysis of the down-regulated differentially expressed genes (DEGs) in NAT10 KD hESCs showed that some genes were enriched in multicellular organismal signaling, second-messenger-mediated signaling and a variety of signaling pathways (Figure 3I and Supplementary Table S6). In addition, GO-BP analysis of the up-regulated DEGs in NAT10 KD hESCs showed significant functional terms related to morphogenesis, development and differentiation, such as kidney morphogenesis, endoderm development and endodermal cell differentiation (Figure 3J and Supplementary Table S7). In summary, the above results indicated that perturbation of NAT10 knock-down impairs self-renewal and promotes early differentiation in hESCs.

Overexpression of NAT10 promotes self-renewal in hESCs

To further investigate whether NAT10 is a regulator of self-renewal in hESCs, we also generated NAT10 OE (overexpression) cell lines by transducing lentiviral particles expressing NAT10 in hESCs. Our data demonstrated that NAT10 OE hESCs maintained the same typical undifferentiated morphology with control hESCs (Figure 4A). ALP staining assays showed that the number of colonies formed by NAT10 OE hESCs was significantly higher than that of control hESCs, suggesting that NAT10 OE hESCs has a significant higher proliferation rate compare to control hESCs (Figure 4B, C). Western blot assay was employed to detect protein expression level of core pluripotent markers in NAT10 OE hESCs and control hESCs, and the result showed that, in comparison with control hESCs, the expressions of OCT4 and SOX2 were significantly increased in NAT10 OE hESCs, while the expression of NANOG was not significantly changed (Figure 4D, E). Together, our data revealed that overexpression of NAT10 is helpful to promote self-renewal in hESCs.

NAT10-deficiency leads to ac^4C loss on target genes

NAT10 is the only known acetyltransferase that catalyzes RNA ac^4C modification (15). To investigate the effect of NAT10 deficiency on ac^4C modification at the transcriptome scale, we performed ac^4C-RIP-seq in the NAT10 KD hESCs, and compared the ac^4C peaks and target genes in the NAT10 KD hESCs with those in the WT hESCs. An overall reduction of ac^4C acetylation peak scores among the ac^4C peaks in the NAT10 KD hESCs could be observed (Figure 5A). Since the ac^4C-RIP-seq may overestimate the prevalence of ac^4C modification on mRNAs, we further introduced $NaCNBH_3$-based chemical ac^4C sequencing (ac^4C-seq) for a rigorous false positive control of ac^4C target genes. WT and NAT10 KD hESCs were investigated with two biological replicates. The result showed that, although the chemical ac^4C sequencing detected many fewer targets than ac^4C-RIP-seq, most sites from chemical ac^4C sequencing could overlap with the acetylation targets detected by ac^4C-RIP-seq. We identified 603 and 559 overlapping ac^4C targets between ac^4C-RIP-seq and chemical ac^4C sequencing in the WT and the NAT10 KD hESCs, respectively (Figure 5B). More importantly, 476 NAT10-responsive ac^4C target genes were finally identified by subtracting acetylation target genes in the WT hESCs by those found in the NAT10 KD hESCs (Figure 5C). This set of 476 NAT10-responsive ac^4C target genes is deemed of high confidence, since it was derived from the consensus results between ac^4C-RIP-seq and orthogonal chemical ac^4C sequencing. The heatmap also showed that the acetylation levels of these NAT10 responsive ac^4C target genes were significantly decreased in the NAT10 KD hESCs (Figure 5D). Next, we performed GO functional enrichment analysis on these NAT10 responsive ac^4C target genes. Interestingly, GO-BP functional enrichment analysis highlighted several functional groups like mitotic nuclear division, DNA

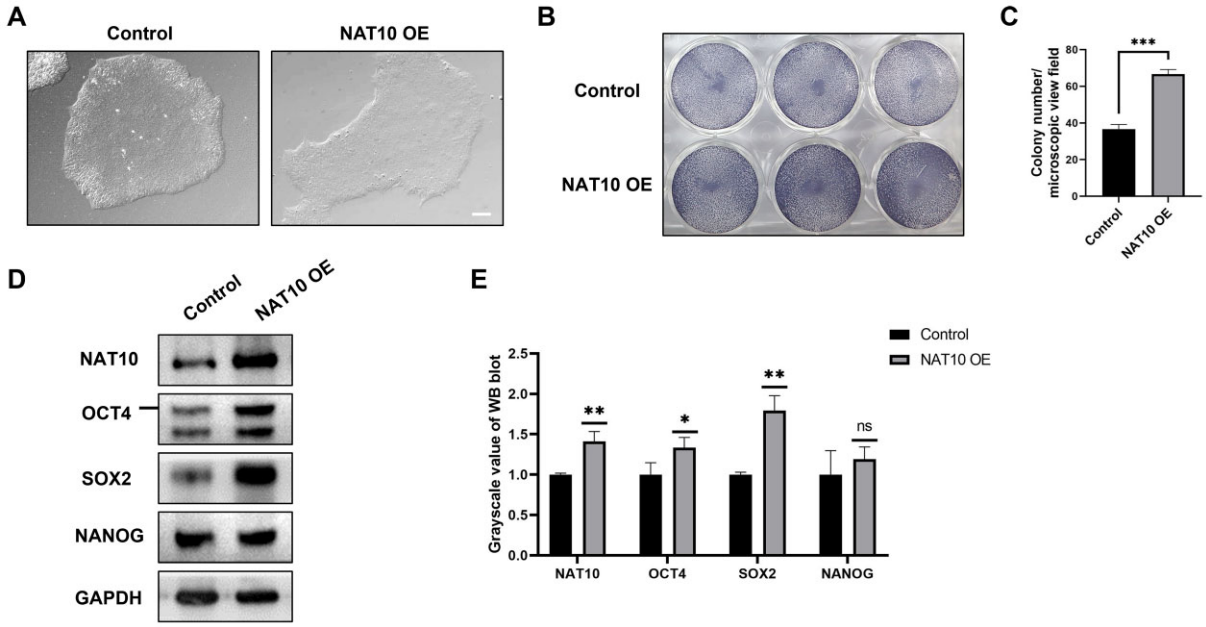


Figure 4. Impact of NAT10 Overexpression on Self-Renewal of hESCs. (**A**) Representative images of morphology of control versus NAT10 overexpression (OE) cells under bright field. Scale bar: 50 μm. (**B**) Alkaline phosphatase (ALP) staining for control and NAT10 OE cells. (**C**) The statistical histogram of number of ALP-positive colony from control versus NAT10 OE cells. *** means $P < 0.001$. (**D**) Representative western blot analysis of expression of NAT10, OCT4, SOX2 and NANOG in control and NAT10 OE cells. OCT4 is the upper band marked by the black line. (**E**) The statistical histogram of grayscale value of western blot. * means $P < 0.05$, ** means $P < 0.01$, ns means not significant. Images in (A), (B) and (D) are representative of three independent experiments. In (C) and (E), error bars represent mean ± SD ($n = 3$ independent experiments).

replication, positive regulation of cell cycle, regulation of mRNA stability, stem cell population maintenance and endoderm development (Figure 5E and Supplementary Table S8). Of note, among these 476 NAT10-responsive ac^{4}C target genes, we observed two core pluripotency regulators, i.e. *OCT4* and *SOX2*.

Furthermore, we assessed the ac^{4}C modification levels in both total RNA and mRNA of WT and NAT10 KD hESCs by using another orthogonal LC–MS/MS approach. The purity of the isolated poly(A)-RNA (mRNA fraction) was verified by RT-qPCR for the detection of contaminations of 18S rRNA and 28S rRNA and leucine/serine tRNAs that are known to have ac^{4}C modifications (13,41), and the results showed no detectable 18S rRNA and 28S rRNA and leucine/serine tRNAs in our purified poly(A)-RNA samples (Supplementary Figure S4A, B). The LC–MS/MS results supported the existence ac^{4}C in both total RNA and the mRNA fraction. Notably, the abundance of ac^{4}C of mRNA was just slightly lower than that of total RNA in WT and NAT10 KD hESCs, indicating that substantial ac^{4}C modifications are distributed on mRNAs in hESCs (Figure 5F). The mean abundance of ac^{4}C modification in WT total RNA and mRNA were 0.43% and 0.36%, followed by that of NAT10 KD at 0.34% and 0.31%, respectively (Figure 5F). Indeed, replicate results suggested that, compared to WT hESCs, the ac^{4}C abundance, either in total RNA or in mRNA, was significantly decreased in NAT10 KD hESCs (Figure 5G, H). Together, these results suggested that the NAT10-catalyzed ac^{4}C acetylation substantially distributed on mRNAs, which are likely to be associated with the cell cycle, mRNA stability and self-renewal of hESCs.

tRNA and rRNA ac^{4}C modification show limited impacts on hESC self-renewal

It is known that tRNA ac^{4}C acetylation also relies on key protein co-factor THUMPD1 (13,16). To further examine and exclude the significant contributions of tRNA ac^{4}C acetylation to hESC self-renewal, we constructed THUMPD1 knockout (KO) hESCs by targeting the *THUMPD1* locus using the CRISPR/Cas9 gene-editing approach (Supplementary Figure S5A). Sanger sequencing of genomic DNA identified that two independent THUMPD1-deficient clones carried the expected deletion (Supplementary Figure S5B). Western blot assay showed that the expression of THUMPD1 was absent in two clones (Supplementary Figure S5C). Morphological images showed that there were no significant differences in cell morphology between THUMPD1 KO hESCs and control hESCs (Figure 6A). ALP staining assay showed that there was no obvious distinction in the number of colonies formed by THUMPD1 KO hESCs in comparison with control hESCs, indicating that THUMPD1 knock-out had no prominent effect on the proliferation rate of hESCs (Figure 6B, C). Next, we employed western blot to detect the protein level of core pluripotent markers OCT4, SOX2 and NANOG in THUMPD1 KO hESCs and control hESCs. Compared with control hESCs, the expressions of OCT4 and NANOG did not have significant change in THUMPD1 KO hESCs and the expression of SOX2 was decreased in only one of the two clones (Figure 6D-E). These results revealed that ac^{4}C modification of tRNA had no significant impact on self-renewal of hESCs.

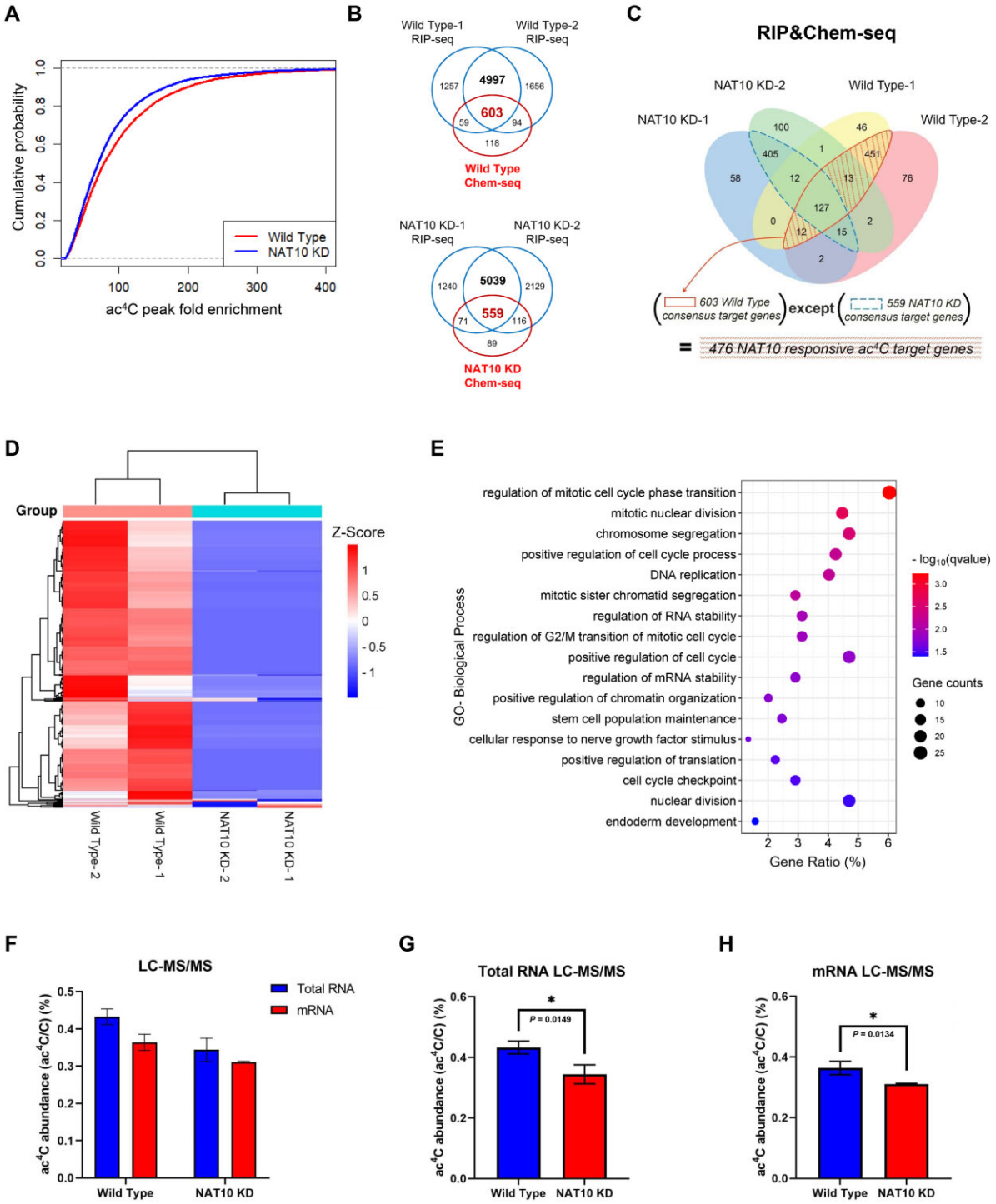


Figure 5. NAT10-catalyzed ac^{4}C Modification Target Genes in hESC Transcriptome. (**A**) Cumulative probability distribution plot comparing ac^{4}C peak fold enrichment score between WT and NAT10 KD cells in ac^{4}C-RIP-seq results. (**B**) Venn diagrams of ac^{4}C target genes in WT and NAT10 KD cells by combing the ac^{4}C-RIP-seq and chemical ac^{4}C sequencing, respectively. Blue circles represent ac^{4}C target genes observed by ac^{4}C-RIP-seq, while red circle represents ac^{4}C sites identified by chemical ac^{4}C sequencing. The red number in the middle represents overlapping ac^{4}C targets between ac^{4}C-RIP-seq and chemical ac^{4}C sequencing. (**C**) Venn diagram of NAT10 responsive ac^{4}C target genes by combing the ac^{4}C peaks detected by ac^{4}C-RIP-seq and ac^{4}C sites detected by chemical ac^{4}C sequencing in WT and NAT10 KD hESCs. The fraction representing the 476 NAT10 responsive ac^{4}C target genes is indicated by shadow. (**D**) Heatmap of fold enrichment scores for 476 NAT10 responsive ac^{4}C target genes among WT and NAT10 KD hESCs. (**E**) Enriched functional terms for Gene GO-BP of the 476 NAT10 responsive ac^{4}C target genes. GO terms with q-value <0.05 were considered statistically significant. The significantly enriched terms that are related to cell cycle, mRNA stability and stem cell population maintenance are shown. (**F**) LC–MS/MS detected the relative ac^{4}C levels (ac^{4}C/C) in total RNA and mRNA from WT and NAT10 KD hESCs. Data are presented as mean ± SD, $n = 3$. (**G**) LC–MS/MS detected the relative ac^{4}C/C abundance in total RNA from WT and NAT10 KD hESCs. (**H**) LC–MS/MS detected the relative ac^{4}C/C abundance in purified mRNA from WT and NAT10 KD hESCs. In (F), (G) and (H), error bars represent mean ± SD ($n = 3$ independent experiments).

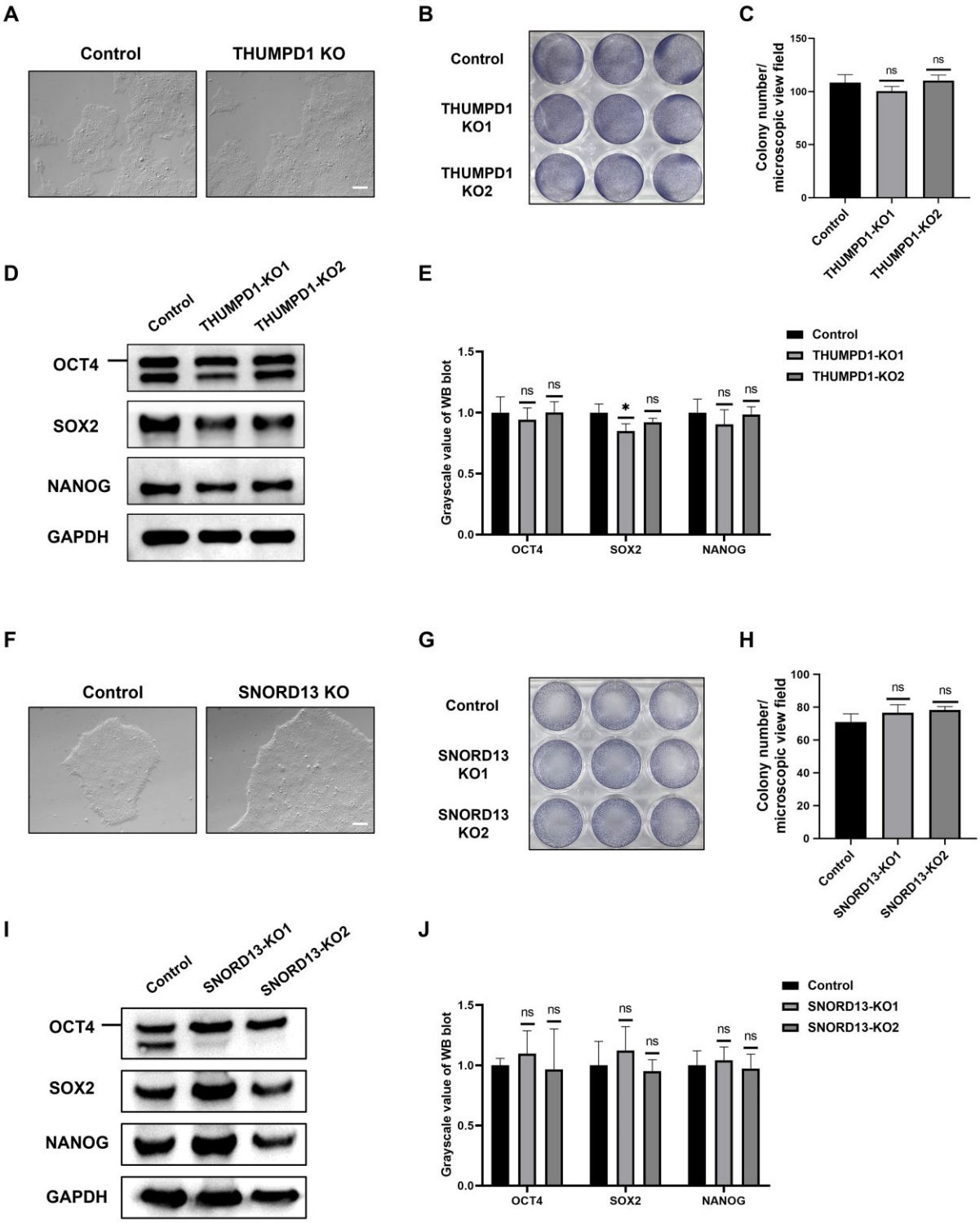


Figure 6. Effect of THUMPD1 and SNORD13 KO on hESC Self-Renewal. (**A**) Representative images of morphology of control versus THUMPD1 knockout (KO) cells under bright field. Scale bar: 50 μm. (**B**) Alkaline phosphatase (ALP) staining for control and THUMPD1 KO cells. (**C**) The statistical histogram of number of ALP-positive colony from control versus THUMPD1 KO cells, ns means not significant. (**D**) Representative western blot analysis of expression of OCT4, SOX2 and NANOG in control and THUMPD1 KO cells. OCT4 is the upper band marked by the black line. (**E**) The statistical histogram of grayscale value of western blot. * means $P < 0.05$, ns means not significant. (**F**) Representative images of morphology of control versus SNORD13 knockout (KO) cells under bright field. Scale bar: 50 μm. (**G**) Alkaline phosphatase (ALP) staining for control and SNORD13 KO cells. (**H**) The statistical histogram of number of ALP-positive colony from control versus SNORD13 KO cells, ns means not significant. (**I**) Representative western blot analysis of expression of OCT4, SOX2 and NANOG in control and SNORD13 KO cells. OCT4 is the upper band marked by the black line. (**J**) The statistical histogram of grayscale value of western blot, ns means not significant. Images in (A), (B), (D), (F), (G) and (I) are representative of three independent experiments. In (C), (E), (H), (J), error bars represent mean ± SD ($n = 3$ independent experiments).

Similarly, small nucleolar RNA SNORD13 is specifically required for rRNA acetylation (17). To gain insight into whether rRNA ac^{4}C acetylation modification plays an important role in hESC self-renewal, we constructed SNORD13 knockout (KO) hESCs by targeting the *SNORD13* locus using the CRISPR/Cas9 gene-editing approach (Supplementary Figure S5D). PCR analysis identified that two independent SNORD13-deficient clones carried the expected deletion (Supplementary Figure S5E, F), which rendered the expression of SNORD13 RNA almost undetectable by RT-qPCR (Supplementary Figure S5G). Morphological image showed that there were no significant differences in cell morphology between SNORD13 KO hESCs and control hESCs (Figure 6F). ALP staining assay showed that there was no obvious difference of colony number formed by SNORD13 KO hESCs in comparison with the control group, indicating that SNORD13 deficiency had no prominent effect on the proliferation rate of hESCs (Figure 6G, H). We also employed western blot to detect the protein level of core pluripotent markers OCT4, SOX2 and NANOG in SNORD13 KO hESCs and control hESCs. Compared with control hESCs, the expressions of OCT4, SOX2 and NANOG had no significant difference in SNORD13 KO hESCs (Figure 6I, J). In conclusion, the above results revealed that perturbation of key co-factors for tRNA/rRNA acetylation did not result in prominent changes in the abovementioned hESC phenotypes, and therefore, rRNA and tRNA ac^{4}C acetylation are unlikely to have significant impact on self-renewal of hESCs.

NAT10-catalyzed ac^{4}C modification targets core pluripotency factor OCT4 and regulates its mRNA stability

As showed in Figure 1D, we observed that ac^{4}C modification would be presented on the majority of pluripotency-related gene transcripts in hESCs through ac^{4}C-RIP-seq. Even further considering rigorous $NaCNBH_3$-based chemical ac^{4}C sequencing, two core pluripotency regulators, *OCT4* and *SOX2*, are highlighted among the 476 high-confident NAT10-responsive ac^{4}C target genes, indicating that these two genes may be noticeable NAT10 responsive ac^{4}C target genes (Figure 7A). Integrative Genomics View (IGV) software was used to interrogate the distribution of ac^{4}C peaks and chemical sequencing-derived ac^{4}C sites across target transcripts of these two core pluripotency genes *OCT4* and *SOX2*. Interestingly, on the *OCT4* transcript, the chemical sequencing-derived ac^{4}C sites in WT hESCs overlapped with its significant ac^{4}C peaks, whereas NAT10 KD hESCs did not have any significant ac^{4}C peak near the chemical sequencing-derived ac^{4}C sites. As for the *SOX2* transcript, WT hESCs had significant ac^{4}C peaks and chemical sequencing-derived ac^{4}C sites, but NAT10 KD hESCs had no chemical sequencing-derived ac^{4}C sites (Figure 7B). Details of all significant chemical sequencing-derived ac^{4}C sites in WT and NAT10 KD hESCs are provided in Supplementary Tables S9-S10. Of note, on *OCT4* transcript, the C-to-T mutation rate of chemical sequencing-derived ac^{4}C sites are 3.7% and 12.3% in WT treatment samples, 3.1% and 4.6% in NAT10 KD treatment samples, both are significant sites (generalized linear mixed-effects model, WT: P = 8.2E-3; NAT10 KD: P = 8.4E-3; Supplementary Figure S6A). Besides, the C-to-T mutation rates in WT treatment samples are also significantly higher than the NAT10 KD counterpart (Chi-squared test, P = 4.8E-2). As for *SOX2* transcript, the C-to-T mutation rates of chemical sequencing-derived ac^{4}C sites are 1.0% and 1.9% in WT treatment samples, indicating a significant ac^{4}C site (generalized linear mixed-effects model, WT: P = 9.8E-3; Supplementary Figure S6B). There is no chemical ac^{4}C site in NAT10 KD hESCs (Supplementary Figure S6B). To further validate the ac^{4}C abundance of *OCT4* and *SOX2*, we performed ac^{4}C-RIP-qPCR assays and used IgG as the isotype control. The results showed that the ac^{4}C abundance of *OCT4* was notably reduced in NAT10 KD cells, while *SOX2* did not have significant change compared to sh control hESCs, which suggests that *SOX2* is less confident NAT10 responsive ac^{4}C target gene (Figure 7C).

Previous study has demonstrated that NAT10-mediated ac^{4}C acetylation modification can exert its functional readouts through the regulation of mRNA stability (mRNA half-life) of its target genes (7). Accordingly, we also performed transcriptome-scale mRNA half-life profiling in NAT10 KD and sh control hESCs, with three parallel experiments per group. Generally, mRNAs in sh control hESCs (with the mean of half-life of 5.78) have significantly longer half-life than the counterparts in NAT10 KD hESCs (with the mean of half-life of 3.47). In other words, perturbation of NAT10 could result in an overall significant decrease in mRNA stability (paired t-test, P = 4.41E-32; Figure 7D). More importantly, when focused on the half-life changes of core pluripotency regulators, we noted that *OCT4* was the one with the most prominent decrease of half-life, with a $\log_2$ fold-change (NAT10 KD vs. sh control) of –1.70, while no prominent decrease was found for *SOX2* (Figure 7E). We also employed the mRNA half-life assay to validate the changes of mRNA stability in response to NAT10 deficiency. The result confirmed that the stability of *OCT4* mRNA is significantly reduced with NAT10 knock-down while the stability of *SOX2* mRNA has no significant change (Figure 7F). At the same time, we confirmed the significant decrease of the protein level of OCT4 in NAT10 KD cells by western blot assay (Figure 7G). In the summary, these results indicate that the core pluripotency-related gene *OCT4* is a unique and noticeable functional target of NAT10-mediated ac^{4}C modification in hESCs, where the mRNA stability and protein expression level of OCT4 are carefully maintained by NAT10.

DISCUSSION

In this study, we have identified that NAT10, the ac^{4}C writer enzyme, positively correlates with pluripotency, where NAT10 overexpression enhances hESC proliferation, while NAT10 knock-down impairs self-renewal and promotes differentiation in hESCs. NAT10-mediated ac^{4}C modifications widely exist on RNAs, and are notably distributed on mRNAs related to self-renewal in hESCs. Mechanistically, mRNAs encoding the core pluripotency regulator *OCT4* is the unique and noticeable NAT10-responsive ac^{4}C modification target gene. NAT10 knock-down significantly reduces stability of *OCT4* mRNA, which leads to hESC

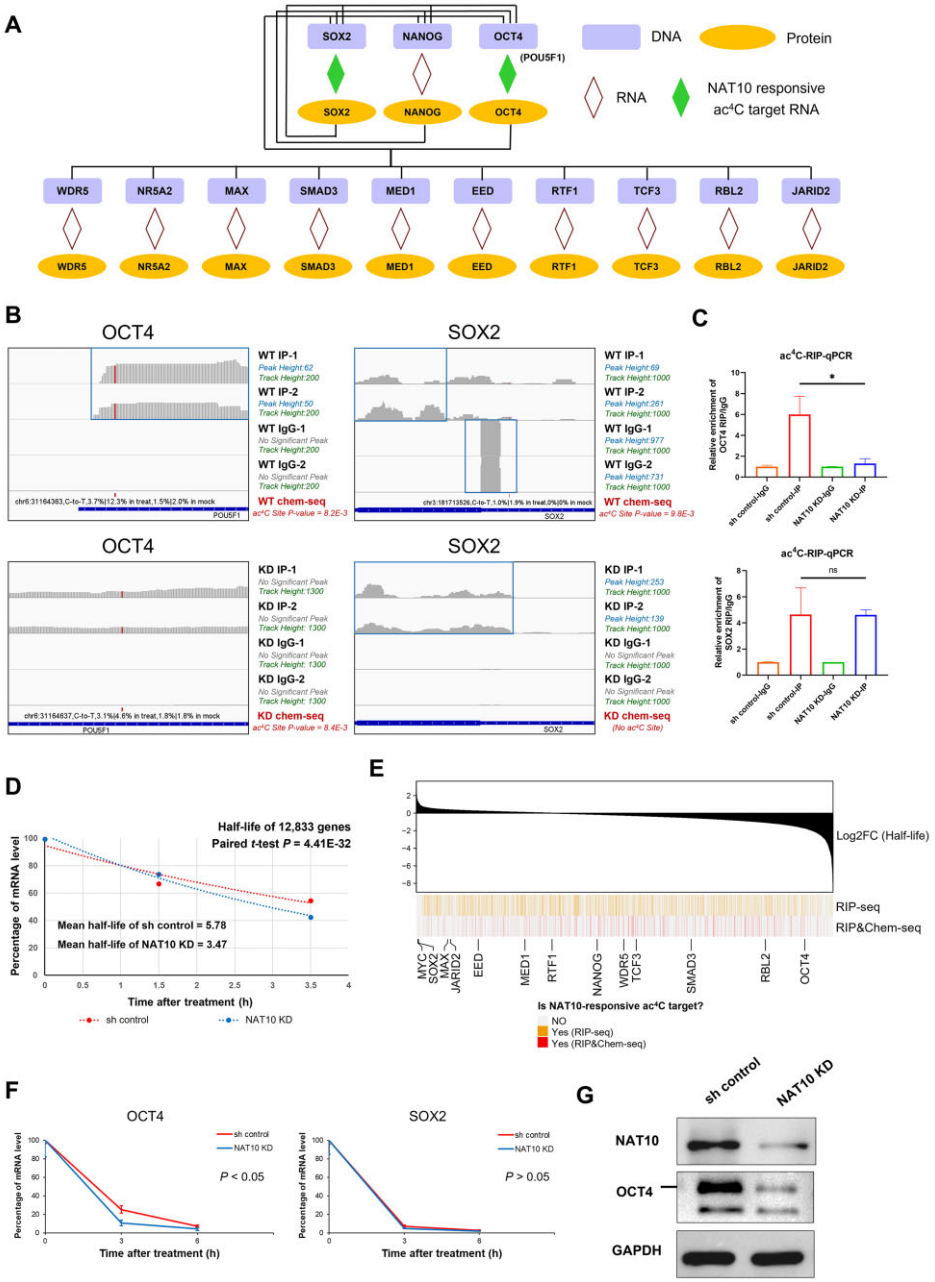


Figure 7. NAT10 regulates target genes in an N^4-acetylcytidine-dependent manner. (**A**) Diagram showing the ac^4C modification for NAT10 responsive ac^4C targets involved in core pluripotency transcriptional network (the network topology is adapted from Young, 2011; permission of reuse has been acquired from Elsevier, with the License Number of 5587090446553). (**B**) Integrative Genomics Viewer (IGV) diagrams displaying read distributions and chemical sequencing-derived ac^4C sites across target transcripts of *OCT4* and *SOX2*. The blue boxes represent significant ac^4C peaks in both replicates of IP or IgG, and the peak height information of significant peaks and the track height information are shown on the right side of each track. It is noticeable that the peak heights were estimated by the coverage function of bedtools, while the track heights were derived from the track height setting of IGV, so these two types of heights are correlated but not fully in the same linear scale. The chemical sequencing-derived ac^4C sites of WT and NAT10 KD samples, if any, are marked as red bars in the ac^4C peaks and below the IP and IgG tracks. For each site, the genomic location and the C-to-T mutation rates in two treatment samples and two mock samples are also specified. The associated *P*-values of chemical ac^4C sites are also shown on the right side of the chemical sequencing tracks. (**C**) ac^4C-RIP-qPCR analysis of alterations in the ac^4C levels of *OCT4* and *SOX2* in NAT10 KD cells versus sh control cells. IgG is used as the isotype control. * means $P < 0.05$, ns means not significant. (**D**) Comparison of the transcriptome-wide half-life profiling results between sh control and NAT10 KD cells. Linear-fitting curve plot of relative mRNA expression level over time is shown. *P*-value is calculated by paired t-test comparing the estimated mRNA half-life of the same gene between two groups. (**E**) Heatmap of half-lives of pluripotency-related gene transcripts. The vertical axis of heatmap represents $\log_2$FC (NAT10 KD vs. sh control) of half-lives of genes. The yellow vertical line below represents the NAT10 responsive ac^4C targets found by ac^4C-RIP-seq, and the red vertical line below represents the NAT10 responsive ac^4C targets by combing ac^4C-RIP-seq and chemical ac^4C sequencing. (**F**) RT-qPCR analysis comparing mRNA decay rate of *OCT4* and *SOX2* in sh control and NAT10 KD cells. (**G**) Representative western blot analysis for expression of NAT10 and OCT4 in sh control and NAT10 KD cells. OCT4 is the upper band marked by the black line. Images in (G) are representative of three independent experiments. In (C) and (F), error bars represent mean ± SD ($n = 3$ independent experiments). It is noteworthy that, for a few samples, the error bars are covered by the corresponding data points because the SD is too small to show.

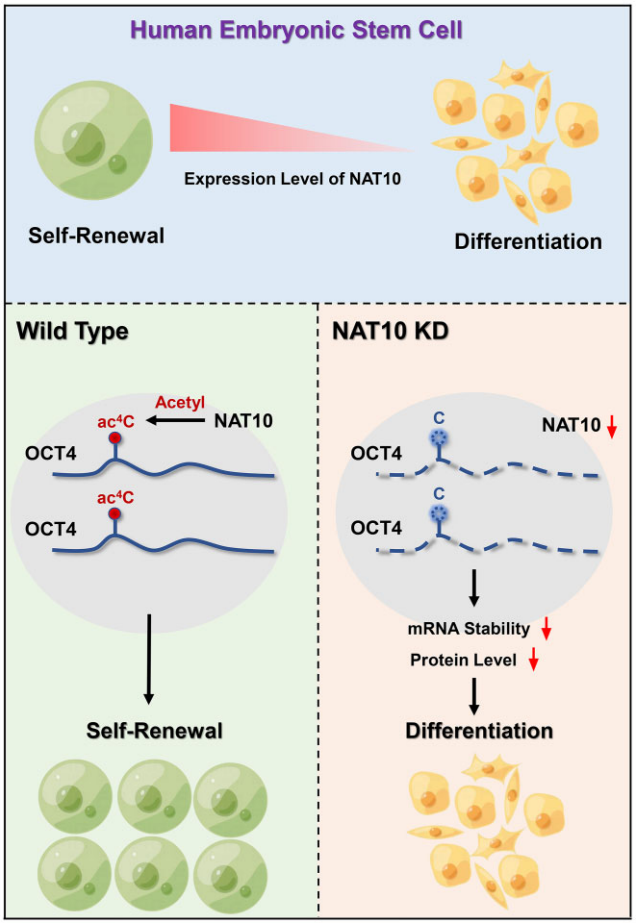


Figure 8. Working model of NAT10-mediated ac^{4}C mRNA modification regulating self-renewal in hESCs. In human embryonic stem cells, NAT10 maintains a high expression so that the hESCs could keep the original pluripotent self-renewal state and the expression level of NAT10 gradually decreased during the process of hESCs from self-renewal to differentiation. Mechanistically, mRNA encoding the core pluripotency regulator *OCT4* modified with ac^{4}C acetylation under the action of RNA acetyltransferase NAT10 in wild-type hESCs, which enhances the stability of *OCT4* mRNA and maintains self-renewal in hESCs. However, NAT10 knock-down results in the loss of ac^{4}C acetylation modification on *OCT4* mRNA, causing the reduction of mRNA stability and protein level of OCT4, and ultimately leads to the differentiation of hESCs.

exit from self-renewal and entering differentiation processes (Figure 8).

Recent study by Arango *et al.* has demonstrated NAT10 as a functional RNA acetyltransferase catalyzes mRNA ac^{4}C modification that stabilizes the acetylated mRNAs (7). In HeLa cells, the level of ac^{4}C modification in mRNA was significantly decreased in the NAT10 knockout cell line. However, the prevalence and functions of ac^{4}C on mRNAs from other, especially non-cancer cell types needs further investigation. Very recently, Chen *et al.* identified the occurrence of ac^{4}C in mRNAs and showed that the modification levels vary among different mouse organs and undergo dynamic changes during spermatogenesis (42). Our results, to our best knowledge, are the first time to demonstrate transcriptome-wide map of ac^{4}C in hESCs. Overall, hESC shares similarity sequence characteristics of ac^{4}C sites with HeLa cells found by Arango *et al.* (7), where ac^{4}C peaks are strongly enriched for cytidine specifically within wobble sites and the repetitive CNN (denoted as CXX by Arango *et al.*) motif. Besides, coding sequence (CDS) is the significantly enriched mRNA partition for ac^{4}C in both HeLa cells and hESCs, although the detailed topologies along mRNAs would be different and a considerable fraction of ac^{4}C acetylation peaks also enriched in 3′-UTRs in hESCs. Nonetheless, ac^{4}C target genes were varied between two cell types. Notably, five pluripotency-related genes *RTF1*, *JARID2*, *WDR5*, *EED* and *TCF3* were also observed to have ac^{4}C modification in HeLa cells, while six pluripotency-related genes *OCT4*, *SOX2*, *NANOG*, *RBL2*, *MED1* and *MAX* were only observed in hESCs, highlighting the specific regulatory targets of ac^{4}C in hESCs.

We are also aware of the controversies about the existence of mRNA ac^{4}C modification in human cells. In 2020, Sas-Chen *et al.* developed ac^{4}C-seq, a chemical genomic method for the transcriptome-wide mapping of ac^{4}C at single-nucleotide resolution. In canonical culture conditions, they detected tRNA and rRNA ac^{4}C sites but not mRNA ac^{4}C sites in HEK293T and HeLa cells (43). However, this does not exclude the presence of mRNA ac^{4}C in other cell types or conditions. In fact, in the same study, Sas-Chen *et al.* also showed that overexpression of NAT10 results in presence of substantial ac^{4}C modification on mRNAs (43). We noted that hESCs are in a similar status with respect to high NAT10 activity. First, the time-course embryonic gene expression atlas has highlighted the high expression level of NAT10 during early embryonic development, especially at the epiblast stage which matches the stage of primed hESCs used in this study (Figure 2A). Moreover, we also observed a marked down-regulation of NAT10 during early differentiation in hESCs (Figure 2B-J). Besides, according to previous LC–MS/MS results in HeLa cells (7), the ac^{4}C modification stoichiometry in mRNA (poly(A)-RNA) fraction is obviously lower than that in total RNA. But our LC–MS/MS results in hESCs has demonstrated largely comparable ac^{4}C modification stoichiometries between mRNAs and total RNAs, after excluding the possible rRNA/tRNA contaminations in mRNA fraction (Figure 5F–H and Supplementary Figure S4). Therefore, it is not surprising that hESCs, showing high NAT10 activity, harbor multiple functional mRNA targets for ac^{4}C modification. Notably, in hESCs, even rigorous chemical ac^{4}C sequencing could detect hundreds of high-confident ac^{4}C modification target genes that were also supported by antibody-based ac^{4}C-RIP-seq approach (Figure 5B). By combing ac^{4}C sites from chemical ac^{4}C sequencing and ac^{4}C peaks from ac^{4}C-RIP-seq, 476 high-confident NAT10-responsive ac^{4}C target genes were finally identified (Figure 5C). These genes tend to be associated with pathways related to cell cycle, mRNA stability and maintenance of stem cell self-renewal (Figure 5E). Among these genes, we also observed two core pluripotency regulators, *SOX2* and *OCT4*. These results indicate the existence of functional regulations by NAT10-mediated mRNA ac^{4}C modification.

Accordingly, we also evaluated the expression and the associated phenotypes of ac^{4}C writer enzyme NAT10 during the proliferation and differentiation of hESCs. We firstly found that the NAT10 expression profile was positively

correlated with pluripotent status, suggesting the functional role of NAT10 in hESCs self-renewal. We then reduced ac^{4}C level through the knock-down of NAT10 and observed attenuation in pluripotent gene expressions and proliferation that eventually leads to the induction of differentiation. We also examined the contribution of tRNA/rRNA ac^{4}C acetylation co-factors in THUMPD1 (a protein factor specifically required for tRNA acetylation) and SNORD13 (a snoRNA specifically required for rRNA acetylation) KO hESCs, and found no significant impact on hESC self-renewal (Figure 6). These results suggest that mRNA ac^{4}C acetylation, but unlikely the canonical tRNA/rRNA ac^{4}C acetylation, are responsible for the NAT10-medidated control of hESC phenotype. Although NAT10 is not a well-defined pluripotency marker like OCT4, we observed that elevated NAT10 expression appears to be a novel characteristic of pluripotent ESCs. To further validate our hypothesis, we overexpressed NAT10 in hESCs and observed that overexpression of NAT10 is indeed helpful to promote self-renewal in hESCs (Figure 4).

Human transcriptome is under the extensive regulations of various types of RNA modification, where m^{6}A methylation is one of the most prevalent modifications in human mRNAs and shows substantial impact on ESC fate decision (3,5,44). Batista *et al.* revealed that the loss of m^{6}A in Nanog transcripts, which delayed Nanog mRNA decay, and impaired mouse ESC exit from self-renewal toward differentiation. In both human and mouse ESCs, the pluripotent state is largely governed by the core transcription factors OCT4, SOX2, and NANOG (45,46). In our study, *OCT4* and *SOX2* were identified as possible NAT10-responsive ac^{4}C target genes by combing the ac^{4}C-RIP-seq and chemical ac^{4}C sequencing, but ac^{4}C-RIP-qPCR verified the reduced ac^{4}C abundance on *OCT4* but not *SOX2* (Figure 7A-C). High throughput mRNA half-life profiling, mRNA stability assay and western blot also suggest that the mRNA stability and protein expression of *OCT4*, but not *SOX2*, were de-regulated in response to NAT10 deficiency (Figure 7D-G). With these validation experiments, we demonstrated that *OCT4* is a key NAT10-responsive ac^{4}C modification target gene. It seems that ac^{4}C modification is involved in regulating self-renewal and differentiation in hESCs by decreasing or increasing *OCT4* mRNA stability, whereas m^{6}A modification is regulated through *NANOG*. Nevertheless, how the two types of RNA modifications synergize to control cell fate transition in hESCs needs to be further investigated.

On the other hand, there are several limitations in current study. First, genetic knock-out of *NAT10* should be a better approach to allow us to examine the loss-of-function phenotypes than knock-down by RNAi. Our data revealed hESCs with complete deletion of NAT10 expression showed complete differentiation morphology and profoundly loss of viability (Supplementary Figure S1). In mice, bi-allelic NAT10 inactivation was lethal at early embryonic development stage, indicating that NAT10 is critical for mouse development (47). Based on above evidence, we speculate that NAT10 may be required for survival in hESCs. Therefore, we chose milder RNAi genetic manipulation to disturb NAT10 expression, investigate NAT10 loss-of-function phenotypes and the underlying mechanisms. We observed residual ac^{4}C level in NAT10 KD cells. One explanation is that there may be other unknown RNA acetyltransferases that have redundant RNA ac^{4}C acetylation activity. The other explanation is that residual (∼20%) NAT10 expression in KD cells is responsible for the remaining ac^{4}C modification.

Second, two main types of pluripotency have been recognized in pluripotent stem cells: naïve cells established from the pre-implantation blastocyst, and primed cells from post-implantation epiblasts (48). Human naive pluripotent stem cells PSCs could be conversed from conventional primed PSCs *in vitro* by culture system optimization (49). So far, seven molecules have been identified as a functional surface marker specific for human naive pluripotency such as ALPPL2, SUSD2, ALPG, DNMT3L, DPPA3, DPPA5, FGF4 (50–52). The hESCs we used in this study demonstrate prime pluripotent status under conventional culture system. We also investigated whether these naïve marker transcripts are modified with ac^{4}C by ac^{4}C-RIP-seq and chemical genomic method for the transcriptome-wide quantitative mapping of ac^{4}C at single-nucleotide resolution. However, ac^{4}C modification of these naïve marker transcripts is not detectable in prime pluripotent status. It is to be determined whether ac^{4}C modification of these naïve marker transcripts emerge in naïve pluripotent stem cells.

Finally, regarding to the ac^{4}C modification, the currently known RNA modification enzyme (Writer) is only the NAT10 protein, but whether there are other modification enzymes in ac^{4}C, as well as its deacetylase (Eraser) and recognition protein (Reader) are not clear. In-depth study of the function and mechanism of ac^{4}C acetylation and the exploration of enzymes involved in its modification can provide more experimental evidence for mechanisms of NAT10 regulating pluripotency. In future, based on the hESC model, we will try to screen the potential ac^{4}C readers and erasers, and investigate their roles and function involved in regulating self-renewal and differentiation.

In summary, the work presented here elucidates a previously unrecognized interconnectivity between the core pluripotent transcriptional network and ac^{4}C RNA epigenetic modification, reveals important insights into the role of NAT10 in the maintenance of hESC self-renewal and found that the core pluripotency regulator *OCT4* is the unique and noticeable NAT10-responsive ac^{4}C modification target gene during hESC self-renewal maintenance. Our study also expanded current understanding of the underlying mechanisms for ESC self-renewal associated with the stem cell epitranscriptome.

DATA AVAILABILITY

The ac^{4}C-RIP-Seq data, chemical ac^{4}C-seq data, gene expression profiling RNA-seq data, and half-life profiling data have been deposited in Gene Expression Omnibus (GEO) database with the Super-Series accession of GSE226752 [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE226752]. The associated ac^{4}C peak tracks are also publicly available in UCSC Genome Browser at https://genome.ucsc.edu/s/soontide/ac4C%2DhESC%2Dpeaks.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the Flow Cytometry Core at National Center for Protein Sciences at Peking University, particularly HY (Huan Yang), for technical help.

FUNDING

National Key R&D Program of China [2022YFC2702704 to Y.L.]; National Natural Science Foundation of China [32070658, 32222020 to Y.Z., 31571517 to Y.L.]; Beijing Municipal Natural Science Foundation [7192091 to Y.L.]; Michigan Medicine-PKUHSC Joint Institute for Translational and Clinical Research [BMU2023JI002 to Y.L.]; Open funding from State Key Laboratory for Reproductive Regulation and Breeding of Grassland Livestock [2021KF0201 to Y.L.]; Open funding from State Key Laboratory of Artificial Microstructure & Mesoscopic Physics (to Y.L.). Funding for open access charge: National Key R&D Program of China [2022YFC2702704 to Y.L.].
Conflict of interest statement. None declared.

REFERENCES

1. Huang,H., Weng,H., Sun,W., Qin,X., Shi,H., Wu,H., Zhao,B.S., Mesquita,A., Liu,C., Yuan,C.L. *et al.* (2018) Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.*, **20**, 285–295.
2. Wang,X., Lu,Z., Gomez,A., Hon,G.C., Yue,Y., Han,D., Fu,Y., Parisien,M., Dai,Q., Jia,G. *et al.* (2014) N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature*, **505**, 117–120.
3. Wang,Y., Li,Y., Toth,J.I., Petroski,M.D., Zhang,Z. and Zhao,J.C. (2014) N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat. Cell Biol.*, **16**, 191–198.
4. Aguilo,F. and Walsh,M.J. (2017) The N(6)-Methyladenosine RNA modification in pluripotency and reprogramming. *Curr. Opin. Genet. Dev.*, **46**, 77–82.
5. Batista,P.J., Molinie,B., Wang,J., Qu,K., Zhang,J., Li,L., Bouley,D.M., Lujan,E., Haddad,B., Daneshvar,K. *et al.* (2014) m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell*, **15**, 707–719.
6. Chen,T., Hao,Y.J., Zhang,Y., Li,M.M., Wang,M., Han,W., Wu,Y., Lv,Y., Hao,J., Wang,L. *et al.* (2015) m(6)A RNA methylation is regulated by microRNAs and promotes reprogramming to pluripotency. *Cell Stem Cell*, **16**, 289–301.
7. Arango,D., Sturgill,D., Alhusaini,N., Dillman,A.A., Sweet,T.J., Hanson,G., Hosogane,M., Sinclair,W.R., Nanan,K.K., Mandler,M.D. *et al.* (2018) Acetylation of cytidine in mRNA promotes translation efficiency. *Cell*, **175**, 1872–1886.
8. Zachau,H.G., Dütting,D. and Feldmann,H. (1966) The structures of two serine transfer ribonucleic acids. *Hoppe Seylers Z Physiol Chem*, **347**, 212–235.
9. Kawai,G., Hashizume,T., Miyazawa,T., McCloskey,J.A. and Yokoyama,S. (1989) Conformational characteristics of 4-acetylcytidine found in tRNA. *Nucleic. Acids. Symp. Ser.*, **21**, 61–62.
10. Kumbhar,B.V., Kamble,A.D. and Sonawane,K.D. (2013) Conformational preferences of modified nucleoside N(4)-acetylcytidine, ac4C occur at "wobble" 34th position in the anticodon loop of tRNA. *Cell Biochem. Biophys.*, **66**, 797–816.
11. Orita,I., Futatsuishi,R., Adachi,K., Ohira,T., Kaneko,A., Minowa,K., Suzuki,M., Tamura,T., Nakamura,S., Imanaka,T. *et al.* (2019) Random mutagenesis of a hyperthermophilic archaeon identified tRNA modifications associated with cellular hyperthermotolerance. *Nucleic Acids Res.*, **47**, 1964–1976.
12. Thomas,G., Gordon,J. and Rogg,H. (1978) N4-Acetylcytidine. A previously unidentified labile component of the small subunit of eukaryotic ribosomes. *J. Biol. Chem.*, **253**, 1101–1105.
13. Sharma,S., Langhendries,J.L., Watzinger,P., Kötter,P., Entian,K.D. and Lafontaine,D.L. (2015) Yeast Kre33 and human NAT10 are conserved 18S rRNA cytosine acetyltransferases that modify tRNAs assisted by the adaptor Tan1/THUMPD1. *Nucleic Acids Res.*, **43**, 2242–2258.
14. Tardu,M., Jones,J.D., Kennedy,R.T., Lin,Q. and Koutmou,K.S. (2019) Identification and quantification of modified nucleosides in saccharomyces cerevisiae mRNAs. *ACS Chem. Biol.*, **14**, 1403–1409.
15. Ito,S., Horikawa,S., Suzuki,T., Kawauchi,H., Tanaka,Y., Suzuki,T. and Suzuki,T. (2014) Human NAT10 is an ATP-dependent RNA acetyltransferase responsible for N4-acetylcytidine formation in 18 S ribosomal RNA (rRNA). *J. Biol. Chem.*, **289**, 35724–35730.
16. Johansson,M.J. and Byström,A.S. (2004) The Saccharomyces cerevisiae TAN1 gene is required for N4-acetylcytidine formation in tRNA. *RNA*, **10**, 712–719.
17. Bortolin-Cavaillé,M.L., Quillien,A., Thalalla Gamage,S., Thomas,J.M., Sas-Chen,A., Sharma,S., Plisson-Chastang,C., Vandel,L., Blader,P., Lafontaine,D.L.J. *et al.* (2022) Probing small ribosomal subunit RNA helix 45 acetylation across eukaryotic evolution. *Nucleic Acids Res.*, **50**, 6284–6299.
18. Jin,G., Xu,M., Zou,M. and Duan,S. (2020) The Processing, gene regulation, biological functions, and clinical relevance of N4-acetylcytidine on RNA: a systematic review. *Mol Ther Nucleic Acids*, **20**, 13–24.
19. Wang,K., Zhou,L.Y., Liu,F., Lin,L., Ju,J., Tian,P.C., Liu,C.Y., Li,X.M., Chen,X.Z., Wang,T. *et al.* (2022) PIWI-interacting RNA HAAPIR regulates cardiomyocyte death after myocardial infarction by promoting NAT10-mediated ac(4) C acetylation of Tfec mRNA. *Adv. Sci. (Weinh.)*, **9**, e2106058.
20. Reubinoff,B.E., Pera,M.F., Fong,C.Y., Trounson,A. and Bongso,A. (2000) Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. *Nat. Biotechnol.*, **18**, 399–404.
21. Broly,M., Polevoda,B.V., Awayda,K.M., Tong,N., Lentini,J., Besnard,T., Deb,W., O'Rourke,D., Baptista,J., Ellard,S. *et al.* (2022) THUMPD1 bi-allelic variants cause loss of tRNA acetylation and a syndromic neurodevelopmental disorder. *Am. J. Hum. Genet.*, **109**, 587–600.
22. Chen,S., Zhou,Y., Chen,Y. and Gu,J. (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, **34**, i884–i890.
23. Kim,D., Langmead,B. and Salzberg,S.L. (2015) HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods*, **12**, 357–360.
24. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G. and Durbin,R. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
25. Meng,J., Cui,X., Rao,M.K., Chen,Y. and Huang,Y. (2013) Exome-based analysis for RNA epigenome sequencing data. *Bioinformatics*, **29**, 1565–1567.
26. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
27. Heinz,S., Benner,C., Spann,N., Bertolino,E., Lin,Y.C., Laslo,P., Cheng,J.X., Murre,C., Singh,H. and Glass,C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, **38**, 576–589.
28. Cui,X., Wei,Z., Zhang,L., Liu,H., Sun,L., Zhang,S.W., Huang,Y. and Meng,J. (2016) Guitar: an R/Bioconductor package for gene annotation guided transcriptomic analysis of RNA-related genomic features. *Biomed. Res. Int.*, **2016**, 8367534.
29. Yu,G., Wang,L.G. and He,Q.Y. (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, **31**, 2382–2383.
30. Thorvaldsdóttir,H., Robinson,J.T. and Mesirov,J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*, **14**, 178–192.
31. Storey,J.D. (2003) The positive false discovery rate: a Bayesian interpretation and the q-value. *Ann. Stat.*, **31**, 2013–2035.
32. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.

33. Piechotta,M., Wyler,E., Ohler,U., Landthaler,M. and Dieterich,C. (2017) JACUSA: site-specific identification of RNA editing events from replicate sequencing data. *BMC Bioinf.*, **18**, 7.
34. Thalalla Gamage,S., Sas-Chen,A., Schwartz,S. and Meier,J.L. (2021) Quantitative nucleotide resolution profiling of RNA cytidine acetylation by ac4C-seq. *Nat. Protoc.*, **16**, 2286–2307.
35. Liao,Y., Smyth,G.K. and Shi,W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.
36. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
37. Gu,Z., Eils,R. and Schlesner,M. (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, **32**, 2847–2849.
38. Dunn,S.J., Martello,G., Yordanov,B., Emmott,S. and Smith,A.G. (2014) Defining an essential transcription factor program for naïve pluripotency. *Science*, **344**, 1156–1160.
39. Young,R.A. (2011) Control of the embryonic stem cell state. *Cell*, **144**, 940–954.
40. Tyser,R.C.V., Mahammadov,E., Nakanoh,S., Vallier,L., Scialdone,A. and Srinivas,S. (2021) Single-cell transcriptomic characterization of a gastrulating human embryo. *Nature*, **600**, 285–289.
41. Boccaletto,P., Stefaniak,F., Ray,A., Cappannini,A., Mukherjee,S., Purta,E., Kurkowska,M., Shirvanizadeh,N., Destefanis,E., Groza,P. *et al.* (2022) MODOMICS: a database of RNA modification pathways. 2021 update. *Nucleic Acids Res.*, **50**, D231–D235.
42. Chen,L., Wang,W.J., Liu,Q., Wu,Y.K., Wu,Y.W., Jiang,Y., Liao,X.Q., Huang,F., Li,Y., Shen,L. *et al.* (2022) NAT10-mediated N4-acetylcytidine modification is required for meiosis entry and progression in male germ cells. *Nucleic Acids Res.*, **50**, 10896–10913.
43. Sas-Chen,A., Thomas,J.M., Matzov,D., Taoka,M., Nance,K.D., Nir,R., Bryson,K.M., Shachar,R., Liman,G.L.S., Burkhart,B.W. *et al.* (2020) Dynamic RNA acetylation revealed by quantitative cross-evolutionary mapping. *Nature*, **583**, 638–643.
44. Geula,S., Moshitch-Moshkovitz,S., Dominissini,D., Mansour,A.A., Kol,N., Salmon-Divon,M., Hershkovitz,V., Peer,E., Mor,N., Manor,Y.S. *et al.* (2015) Stem cells. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science*, **347**, 1002–1006.
45. Niwa,H. (2007) How is pluripotency determined and maintained?*Development*, **134**, 635–646.
46. Silva,J. and Smith,A. (2008) Capturing pluripotency. *Cell*, **132**, 532–536.
47. Balmus,G., Larrieu,D., Barros,A.C., Collins,C., Abrudan,M., Demir,M., Geisler,N.J., Lelliott,C.J., White,J.K., Karp,N.A. *et al.* (2018) Targeting of NAT10 enhances healthspan in a mouse model of human accelerated aging syndrome. *Nat. Commun.*, **9**, 1700.
48. Dundes,C.E. and Loh,K.M. (2020) Bridging naïve and primed pluripotency. *Nat. Cell Biol.*, **22**, 513–515.
49. Theunissen,T.W., Powell,B.E., Wang,H., Mitalipova,M., Faddah,D.A., Reddy,J., Fan,Z.P., Maetzel,D., Ganz,K., Shi,L. *et al.* (2014) Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell*, **15**, 471–487.
50. Bi,Y., Tu,Z., Zhang,Y., Yang,P., Guo,M., Zhu,X., Zhao,C., Zhou,J., Wang,H., Wang,Y. *et al.* (2020) Identification of ALPPL2 as a naive pluripotent state-specific surface protein essential for human naive pluripotency regulation. *Cell Rep.*, **30**, 3917–3931.
51. Bredenkamp,N., Stirparo,G.G., Nichols,J., Smith,A. and Guo,G. (2019) The Cell-surface marker sushi containing domain 2 facilitates establishment of human naive pluripotent stem cells. *Stem Cell Rep.*, **12**, 1212–1222.
52. Collier,A.J., Panula,S.P., Schell,J.P., Chovanec,P., Plaza Reyes,A., Petropoulos,S., Corcoran,A.E., Walker,R., Douagi,I., Lanner,F. *et al.* (2017) Comprehensive cell surface protein profiling identifies specific markers of human naive and primed pluripotent states. *Cell Stem Cell*, **20**, 874–890.