STEM CELLS

N-acetyltransferase NAT10 controls cell fates via connecting mRNA cytidine acetylation to chromatin signaling

Zhensheng Hu[1]†, Yunkun Lu[1]†, Jie Cao[2,3]†, Lianyu Lin[1]†, Xi Chen[1], Ziyu Zhou[1], Jiaqi Pu[1,4], Guo Chen[1], Xiaojie Ma[1], Qian Deng[1], Yan Jin[1], Liling Jiang[1], Yuhan Li[1], Tengwei Li[2], Jianzhao Liu[2,3]*, Saiyong Zhu[1]*

Cell fate transition involves dynamic changes of gene regulatory network and chromatin landscape, requiring multiple levels of regulation, yet the cross-talk between epitranscriptomic modification and chromatin signaling remains largely unknown. Here, we uncover that suppression of *N*-acetyltransferase 10 (NAT10), the writer for mRNA N^4-acetylcytidine (ac^4C) modification, can notably affect human embryonic stem cell (hESC) lineage differentiation and pluripotent reprogramming. With integrative analysis, we identify that NAT10-mediated ac^4C modification regulates the protein levels of a subset of its targets, which are strongly enriched for fate-instructive chromatin regulators, and among them, histone chaperone ANP32B is experimentally verified and functionally relevant. Furthermore, NAT10-ac^4C-ANP32B axis can modulate the chromatin landscape of their downstream genes (e.g., key regulators of Wnt and TGFβ pathways). Collectively, we show that NAT10 is an essential regulator of cellular plasticity, and its catalyzed mRNA cytidine acetylation represents a critical layer of epitranscriptomic modulation and uncover a previously unrecognized, direct cross-talk between epitranscriptomic modification and chromatin signaling during cell fate transitions.

INTRODUCTION

Understanding cellular plasticity is critical in biology. Human pluripotent stem cells provide a very important and tractable system for extensively studying cell fate transitions (*1*–*4*). Cell fate transitions prompt widespread changes in gene regulatory network and chromatin landscape. Although many efforts have been devoted to understanding the underlying mechanisms directly at epigenetic and transcriptional levels, more investigations are still urgently required at posttranscriptional levels. Particularly, compared to the widely studied epigenetic regulations on DNAs and histones, epitranscriptomic regulations on RNAs for controlling cell fates have only become appreciated in recent years (*5*–*10*). Studying these different layers of regulation and their cross-talk will provide better understanding of cell fate transitions and can be harnessed for regenerative medicine.

Acetylation of mRNAs, named N^4-acetylcytidine (ac^4C), has recently been identified as an epitranscriptomic modification. Several studies reported the existence of ac^4C modifications in mRNAs from HeLa and human embryonic kidney (HEK) 293T cells and demonstrated that ac^4C can regulate mRNA stability and translation efficiency (*11*, *12*), boost HIV-1 gene expression (*13*), and modulate tumorigenesis (*14*, *15*). *N*-acetyltransferase 10 (NAT10) is the singular human enzyme to have both acetyltransferase and RNA binding activities, and more recently, NAT10 has been identified as the only known writer enzyme for ac^4C modification on mRNAs (*11*, *12*). NAT10 is highly overexpressed in multiple tumor types, including hepatocellular carcinoma, colorectal cancer, and melanoma (*16*–*18*). NAT10 is also related to Hutchinson-Gilford progeria syndrome (*19*, *20*), and targeting of *NAT10* can enhance healthspan in a mouse model of human accelerated aging syndrome (*21*). So far, the functions of NAT10 and mRNA ac^4C modification in PSCs, or during cell fate transitions, remain largely unexplored.

Here, we demonstrate that knockdown of *NAT10* can significantly impair human embryonic stem cell (hESC) proliferation, affect hESC differentiation, and inhibit pluripotent reprogramming. In addition, we examine the ac^4C levels of mRNAs by liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) and find that mRNA ac^4C levels are notably reduced after *NAT10* suppression. The acetylated RNA immunoprecipitation followed by sequencing (acRIP-seq) results demonstrate that NAT10 mediates the ac^4C modifications on many downstream target transcripts, which are enriched for epigenetic regulator genes, such as *ANP32B*. Functionally, ANP32B and NAT10 are relevant. In addition, NAT10-ac^4C-ANP32B axis can modulate chromatin landscape of their downstream genes (e.g., *SFRP1* and *NODAL*). Collectively, we identify NAT10 as an important regulator of pluripotent reprogramming and differentiation, demonstrate that RNA cytidine acetylation represents an essential layer of epitranscriptomic regulation, and uncover NAT10-ac^4C-ANP32B axis as a previously unknown, direct link between epitranscriptome and chromatin signaling during cell fate transitions.

RESULTS

NAT10 has a broad role in regulating hESC properties

When we did differentiation and reprogramming experiments, we found that *NAT10* expression was very dynamic during cell fate transitions (fig. S1, A and B). To further determine the roles of

[1]Life Sciences Institute, The Second Affiliated Hospital and School of Medicine, The MOE Key Laboratory of Biosystems Homeostasis and Protection and Zhejiang Provincial Key Laboratory for Cancer Molecular Cell Biology, Zhejiang University, Hangzhou, Zhejiang 310058, China. [2]MOE Key Laboratory of Macromolecular Synthesis and Functionalization, Department of Polymer Science and Engineering, Zhejiang University, Hangzhou 310058, China. [3]Life Sciences Institute, Zhejiang University, Hangzhou 310058, China. [4]The Children's Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China.
*Corresponding author. Email: saiyong@zju.edu.cn (S.Z.); liujz@zju.edu.cn (J.L.)
†These authors contributed equally to this work.

NAT10 in stem cell biology, we applied a short hair RNA (shRNA)–based approach to knock down *NAT10* in hESCs (Fig. 1A). Using Western blotting, we confirmed that the protein level of NAT10 was successfully reduced (fig. S1C). Immunostaining further confirmed the decrease of NAT10 protein in the nucleolus (fig. S1D). Typical pluripotency markers, including OCT4, NANOG, and SOX2, were almost unaffected (fig. S1, E and F). We observed that the colony size was reduced in sh*NAT10* hESCs (Fig. 1B). Growth rates were also decreased (Fig. 1C). In addition, hESCs with *NAT10* suppression resulted in a substantial reduction of DNA and protein syntheses, mimicking a quiescent state (Fig. 1, D and E). Immunoprecipitation and MS further showed that most of the NAT10-interacted proteins were enriched in Gene Ontology (GO) terms of ribosome biogenesis and cell cycle (fig. S1G). Collectively, these results indicated that knockdown of *NAT10* in hESCs induces a quiescent state without affecting the core pluripotent gene regulatory network.

We then applied RNA sequencing (RNA-seq) to evaluate the functions of NAT10. In total, 1872 genes were up-regulated and 1538 genes were down-regulated after *NAT10* knockdown in hESCs ($|\text{fold change}| > 1.5$, adjusted P value < 0.05) (fig. S2A). We did Kyoto Encyclopedia of Genes and Genomes (KEGG) and GO analyses. Up-regulated genes belonged to terms including p53 signaling pathway (e.g., *GADD45A* and *GADD45B*), oxidative phosphorylation (e.g., *COX7A1* and *NDUFA8*), negative regulation of cell cycle (e.g., *CDKN1A* and *CDKN2A*), and apoptosis (e.g., *CASP3* and *CASP10*) (Fig. 1, F and H, and fig. S2B), which explained the phenotype of hESC proliferation delay upon *NAT10* suppression. The down-regulated terms included glycine, serine, and threonine metabolism (e.g., *CTH* and *GCAT*), glycolysis (e.g., *HK2* and *PFKM*), Wnt signaling pathway (e.g, *FZD2* and *CCND1*), and Hippo signaling pathway (e.g., *YAP1*, and *TEAD1*) (Fig. 1, G and I, and fig. S2C), suggesting that *NAT10* is probably involved in regulating cell fate determination in hESCs.

Epigenetic landscape of hESCs is highly correlated with their developmental capacity. In particular, H3K4me3 and H3K27me3 are associated with gene activation and repression, respectively, and bivalency (with both H3K4me3 and H3K27me3 modifications) is a key epigenetic feature of hESCs (*22*), contributing their differentiation potentials. Then, we examined genome-wide H3K4me3 and H3K27me3 modifications after *NAT10* suppression (fig. S2, D to G), and the results showed that *NAT10* knockdown affected chromatin landscape, which is correlative to gene expression. For example, regions with up-regulated H3K4me3 were related to tumor necrosis factor (TNF) signaling and p53 pathway (Fig. 1, J and K). Bivalency was also affected by *NAT10* suppression, with enriched GO terms of signal transduction, regulation of Wnt signaling pathway, and cell differentiation (fig. S2, H and I). Collectively, NAT10 plays important roles in proliferation, signaling, and epigenetics and potentially regulates stem cell differentiation.

NAT10 is an important regulator for lineage differentiation

On the basis of the observation of widespread changes of gene expression and histone modifications after *NAT10* knockdown, we next investigated the developmental potentials of sh*NAT10* hESCs (Fig. 2A). First, embryoid body (EB) results showed that *NAT10* knockdown severely affected normal EB formation, suggesting that NAT10 is important for spontaneous differentiation (Fig. 2B). Second, we applied a directed pancreatic differentiation protocol. At definitive endoderm stage, both shCTR and sh*NAT10* cells could generate SOX17 and FOXA2 double-positive cells (fig. S2, A and B). At later stage, pancreatic differentiation was significantly affected (Fig. 2, C and D). We could observe the induction of pancreatic progenitor markers PDX1 and NKX6.1 in control cells, while these two markers were hardly detected in sh*NAT10* cells (Fig. 2C). Reverse transcription quantitative polymerase chain reaction (RT-qPCR) result further confirmed the reduction of key pancreatic progenitor marker genes, including *PDX1*, *NKX6.1*, *SOX9*, and *HNF6*, in sh*NAT10* cells (Fig. 2D). Last, we evaluated the in vivo developmental potentials of these sh*NAT10* cells by teratoma assay. Control cells could successfully form teratomas containing tissues of three germ layers, while 4 million sh*NAT10* cells injected into immune-deficient SCID Beige mice failed to form teratoma (fig. S3C). Therefore, NAT10 is particularly important for lineage specification.

NAT10 is necessary for pluripotent reprogramming

To further evaluate the roles of NAT10 in cell fate transitions, we resorted to transcription factor–induced pluripotent reprogramming, which is widely used for studying cell plasticity (Fig. 2E). Knockdown of *NAT10* by shRNAs notably blocked the formation of alkaline phosphatase (AP) and OCT4-positive induced pluripotent stem cell (iPSC) colonies (Fig. 2, F and G). Quantitatively, fluorescence-activated cell sorting (FACS) results showed that OCT4-positive cell number was significantly decreased after *NAT10* knockdown (fig. S3D). Using RT-qPCR, we found that sh*NAT10* altered the induction of *CDH1* expression and repression of *CDH2* expression, indicating the unsuccessful mesenchymal-to-epithelial transition (MET) (Fig. 2H). The induction of pluripotency marker genes, such as *OCT4*, *SOX2*, *NANOG*, and *SALL4*, was all impaired by *NAT10* knockdown (Fig. 2I). In addition, we knocked down *NAT10* by single guide RNAs (sgRNAs) and observed similar phenotypes (fig. S3, E to J). Last, we applied a chemical approach using Remodelin that has previously been reported to targeting NAT10 (fig. S3E). Remodelin inhibited the formation of AP and OCT4-positive iPSC colonies, reduced the percentage of OCT4-positive cells, delayed MET, and blocked the induction of pluripotent marker genes (fig. S3, K to Q). Collectively, NAT10 also plays vital roles in cellular reprogramming.

NAT10 regulates the mRNA ac^{4}C levels in hESCs

NAT10 is recently identified and suggested as the singular ac^{4}C writer and the principal source of ac^{4}C modification in HeLa and HEK293T cells (Fig. 3A) (*11*, *12*). However, it was reported that ac^{4}C was present at a very low level in endogenous eukaryotic mRNAs (*23*). To explore the existence of ac^{4}C in normal stem cells, particularly hESCs, we established a stable LC-MS/MS system (fig. S4A). To exclude the contamination from ribosomal RNAs (rRNAs), we enriched polyadenylate [poly(A)] RNAs through oligo(dT) selection and RNA levels of 18*S* rRNA were quantified via RT-qPCR (Fig. 3B). The abundance of ac^{4}C modification in hESC total RNAs and poly(A) RNAs was ~0.26 and ~0.09%, respectively (Fig. 3C). The ac^{4}C level of poly(A) RNAs was higher in hESCs than in fibroblasts or in cells after EB differentiation (fig. S4B). *NAT10* knockdown by shRNAs decreased the ac^{4}C levels in hESC total RNAs by ~55 and ~42% (Fig. 3D). Consistent with the degree of ablation in total RNAs, ac^{4}C levels in poly(A) RNAs from sh*NAT10* cells reduced by about 60% (Fig. 3, E and F). These data demonstrated the existence of ac^{4}C modification in mRNAs of hESCs, and NAT10 indeed regulates the ac^{4}C levels of mRNAs in hESCs.

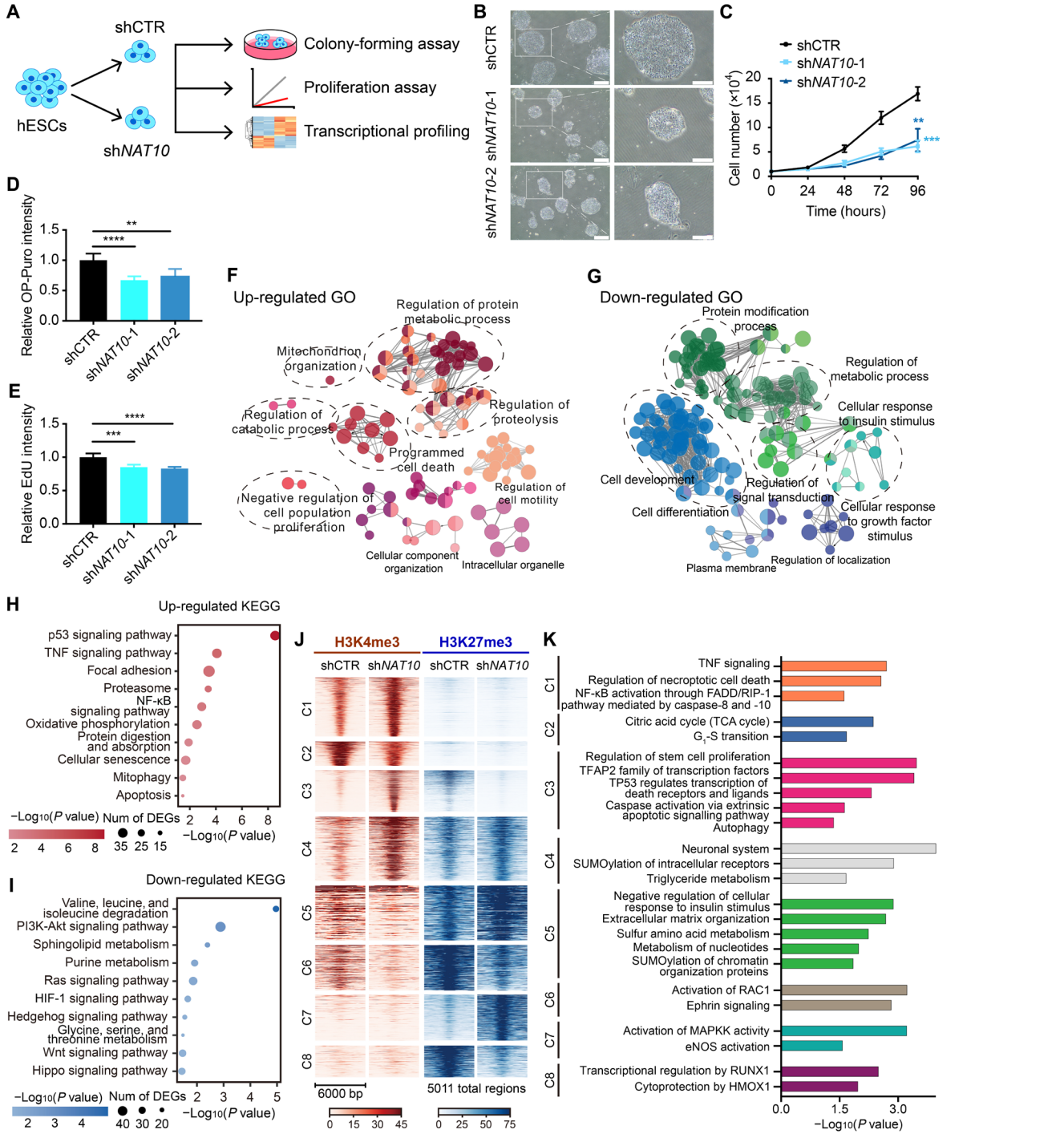

Fig. 1. Characterization of *NAT10* knockdown hESCs. (**A**) Experimental scheme for (B) to (I). (**B**) Light microscopy images of day 6 colony formation (left) and representation of colony morphology (right) for hESCs transduced with shCTR or sh*NAT10*, respectively. Scale bars, 500 μm (left) and 1 mm (right). (**C**) Growth curves of hESCs transduced with shCTR or sh*NAT10*. Data are means ± SD of $N = 3$. (**D**) Measurement of protein synthesis. Data are means ± SD of $N = 6$. (**E**) Measurement of DNA synthesis. Data are means ± SD of $N = 6$. (**F** and **G**) Networks showing GO enrichment results for up-regulated and down-regulated genes. Each dashed circle represents one enriched GO term. GO terms were grouped on the basis of overlapped genes, and the representative GO terms for each group were labeled. (**H** and **I**) Bubble plots showing KEGG enrichment of up-regulated (red) and down-regulated (blue) genes. Gradient colors represent enriched significance, and size of circles represents numbers of differentially expressed genes. (**J**) Heatmaps displaying H3K4me3 and H3K27me3 signals within a 6-kb window of identified differential peaks. The differential peaks are classified into eight separate clusters by *K*-means clustering. (**K**) GO categories of genes associated with differential peaks. $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$. DEGs, differentially expressed genes; PI3K, phosphatidylinositol 3-kinase; HIF-1, hypoxia-inducible factor–1; TNF, tumor necrosis factor; TCA, tricarboxylic acid; FADD, fas-associated death domain; RIP-1, receptor-interacting protein 1; MAPKKK, MAPK kinase kinase; eNOS, endothelial nitric oxide synthase; RUNX-1, Runt-related transcription factor 1; HMOX-1, heme oxygenase 1; TFAP2, transcription factor AP-2.

ac^{4}C modifications on mRNAs are mapped in hESCs

To identify potential targets of NAT10 and investigate whether ac^{4}C of mRNAs plays a role in cell fate determination, we performed acRIP-seq. Compared to the most prevalently studied mRNA m^{6}A, the study of ac^{4}C is much more challenging due to the lack of excellent antibody and relatively low modification level. To overcome these issues, we used ac^{4}CTP to elute acetylated RNA and detected ac^{4}C peaks with high specificity (Fig. 4A). Critically, ac^{4}C(+)-RNA probe was significantly enriched, validating the efficiency of ac^{4}C immunoprecipitation (Fig. 4B). Overall, we obtained 2321 and 1492 ac^{4}C peaks in shCTR and sh*NAT10* hESCs, with 889 of them were overlapped (Fig. 4, C and D, and fig. S5A). *CCDC34* and *POU5F1* are represented for highly enriched ac^{4}C target and nonenriched control, respectively (Fig. 4E). Approximately 75% of ac^{4}C peaks were located at gene coding regions (Fig. 4F). Furthermore, the distribution of the peaks across transcripts displayed a significant enrichment in gene coding regions in vicinity to 3′ untranslated regions (Fig. 4G). The most significant motif of ac^{4}C peaks is "CXUCXUCXUCXU" (Fig. 4H), and we also found other similar motifs in the high-ranking motif enrichment results (fig. S5B).

We found that distributions of the ac^{4}C peaks across chromosomes have almost no correlation with gene distribution, ac^{4}C were relatively stable in targeting genes, and most of the ac^{4}C sites remained existing upon *NAT10* knockdown (fig. S5C). These ac^{4}C-targeting

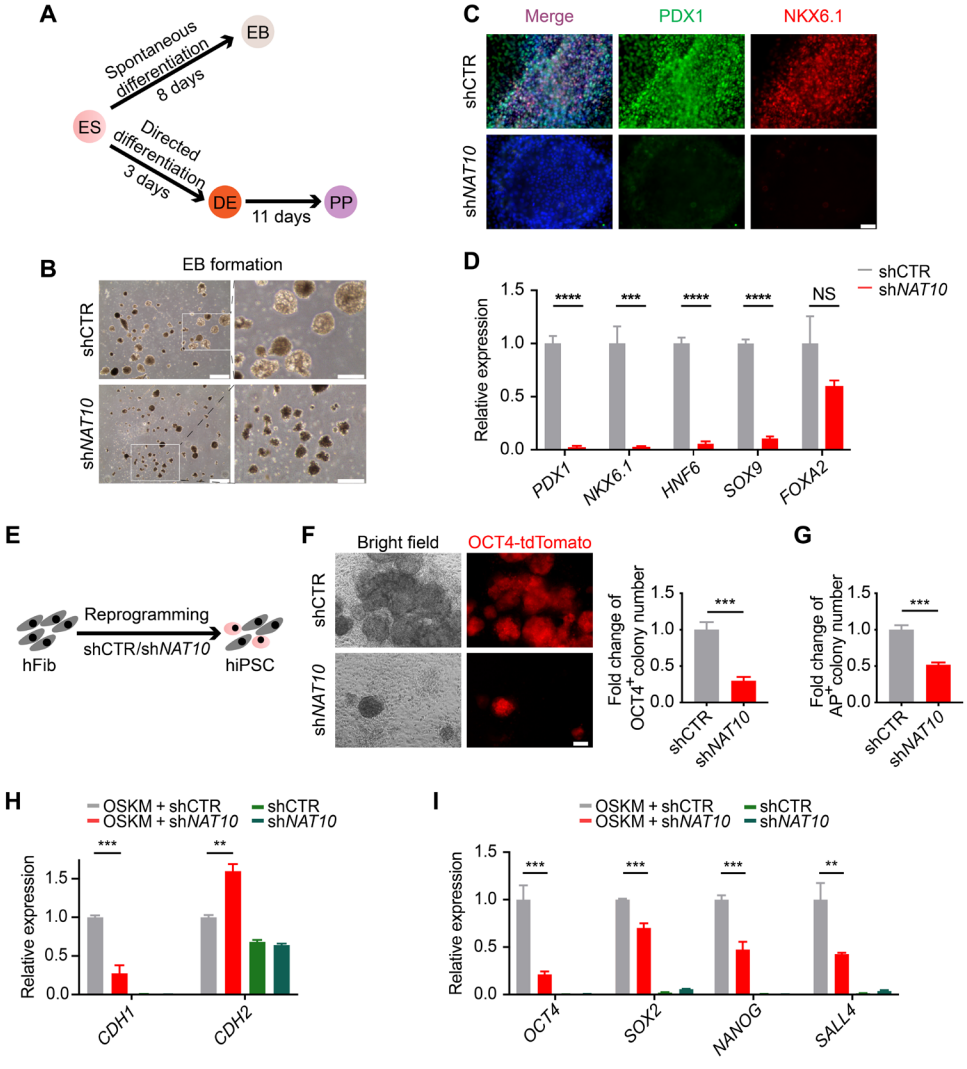

Fig. 2. NAT10 plays essential roles in stem cell differentiation and cellular reprogramming. (**A**) Schematic diagram depicting the strategy of spontaneous and directed differentiation of hESCs. (**B**) Light microscopy images of day 8 EB formation (left) and representation of EB morphology (right). Scale bars, 500 μm (left) and 1 mm (right). (**C**) Immunofluorescence of PDX1 and NKX6.1 at day 14 of pancreatic differentiation. Nuclei were co-stained with Hoechst (blue). Scale bar, 50 μm. (**D**) RT-qPCR analysis of *PDX1*, *NKX6.1*, *HNF6*, *SOX9*, and *FOXA2* at day 14 of pancreatic differentiation. Data are means ± SD of $N = 3$. (**E**) Diagram showing the procedure of *NAT10* knockdown by shRNA. (**F**) Representation of OCT4-positive colony in shCTR cells and sh*NAT10* cells (left) and quantification (right). Data are means ± SD of $N = 3$. Scale bar, 100 μm. (**G**) Quantification of alkaline phosphatase staining in shCTR cells and sh*NAT10* cells. Data are means ± SD of $N = 3$. (**H**) RT-qPCR analysis of MET gene expression in cells treated with OSKM plus shCTR and sh*NAT10* or shCTR and sh*NAT10*. Data are means ± SD of $N = 3$. (**I**) RT-qPCR analysis of pluripotency gene expression in cells treated with OSKM plus shCTR and sh*NAT10* or shCTR and sh*NAT10*. Data are means ± SD of $N = 3$. $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$. hFib, human fibroblast; hiPSC, induced pluripotent stem cell; NS, not significant.

genes were significantly enriched in ribosome, spliceosome, oxidative phosphorylation, and proteasome (Fig. 4I). Previous studies have shown the association between ac^{4}C formation and pre-mRNA retention and splicing complex (*24*). To examine the influences of *NAT10* knockdown on pre-mRNA splicing, we calculated percent spliced in (PSI) and percent intron retention (PIR) for shCTR and sh*NAT10*, separately. We observed that PSI values were slightly reduced in sh*NAT10*, indicating modest effects of NAT10 on exon inclusion/exclusion (Fig. 4J). Notably, PIR values were notably increased in sh*NAT10*, strongly implying the reduced splicing efficiency upon *NAT10* knockdown (Fig. 4K). Collectively, we found that NAT10 could affect pre-mRNA splicing by targeting spliceosomal genes.

Next, we found that the ac^{4}C abundance in sh*NAT10* hESCs was substantially lower (Fig. 4L), which is consistent with the results of LC-MS/MS that *NAT10* suppression could substantially decrease ac^{4}C modification level (Fig. 3, E and F). Moreover, differential peak analysis revealed 143 down-regulated peaks and one up-regulated peak (Fig. 4M). Functional enrichment analysis for the genes with down-regulated peaks showed significant enrichment in spliceosome and ribosome biogenesis (fig. S5, D and E). Collectively, these results demonstrated that the ac^{4}C levels of mRNAs were significantly reduced upon *NAT10* knockdown, which subsequently affect pre-mRNA splicing and translation. However, the abovementioned broad effects of NAT10 on cell fate transitions remain largely unexplained.

Chromatin regulators are key downstream targets of NAT10

Recent studies have suggested the association of ac^{4}C in mRNA coding sequences (CDSs) with mRNA stability and translation efficiency (*11*, *12*). To identify the key downstream targets of NAT10 that can contribute its broad effects on cell fate transitions, we decided to perform large-scale quantitative proteomic analysis. Overall, the changes of protein and gene expression patterns upon *NAT10*

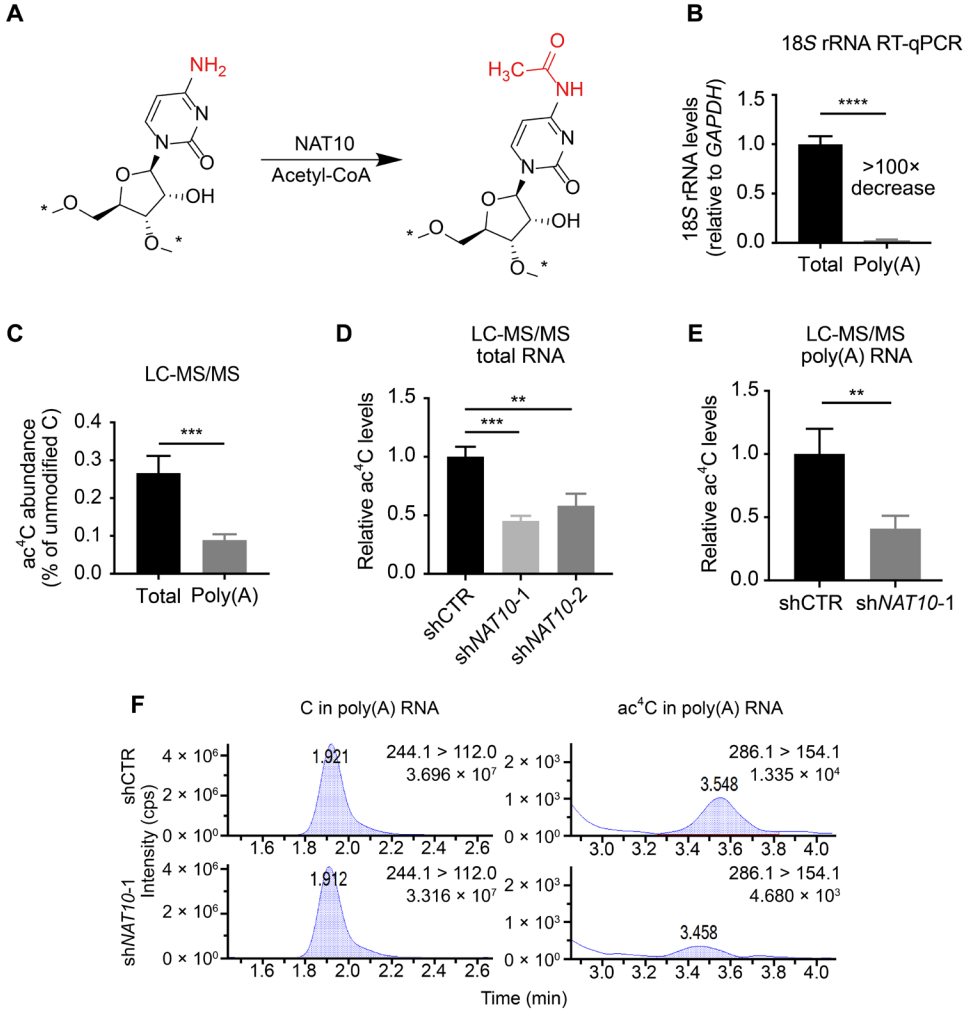

Fig. 3. ac^{4}C detection in hESC mRNAs. (**A**) NAT10 catalyzes the formation of N^4-acetylcytidine (ac^{4}C) modification. (**B**) Determination of poly(A) RNA purity through 18*S* rRNA RT-qPCR. Data are means ± SD of $N = 3$. (**C**) LC-MS/MS of total and poly(A) RNAs from hESCs. Data are means ± SD of $N = 4$. (**D**) Total RNAs were digested to mononucleosides and analyzed by LC-MS/MS. Data are means ± SD of $N = 3$. (**E**) Relative quantification of ac^{4}C detection in poly(A) RNA from hESCs transduced with shCTR or sh*NAT10*. Data are means ± SD of $N = 4$. (**F**) Chromatograms of representative LC-MS/MS performed in poly(A) RNAs from hESCs transduced with shCTR or sh*NAT10*. $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$.

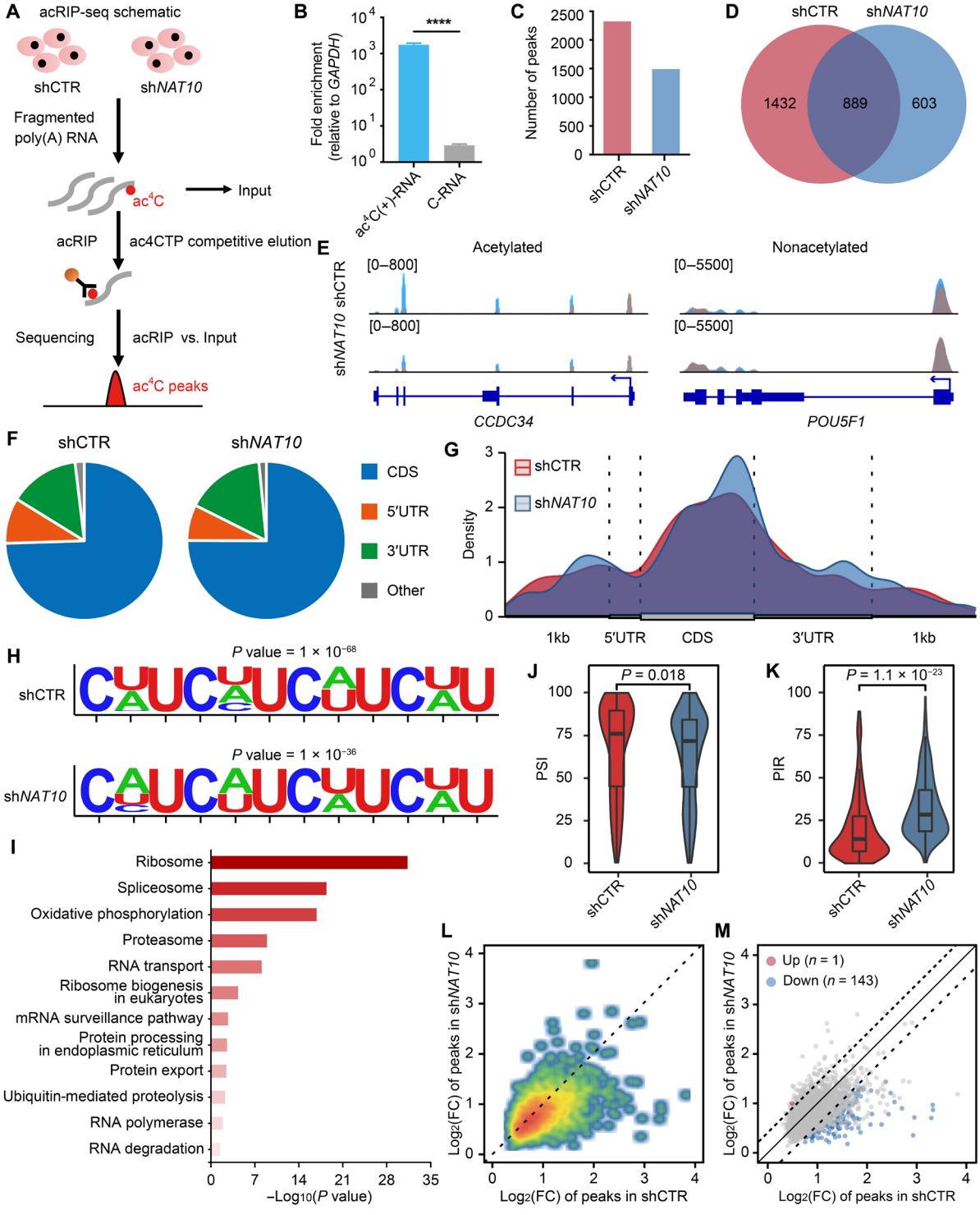

Fig. 4. Mapping of mRNA ac^{4}C in hESCs by acRIP-seq. (**A**) Schematic of acRIP-seq. (**B**) ac^{4}C(+) or C-RNA probes were spiked into poly(A) RNA followed by acRIP-RT-qPCR. Data are means ± SD of $N = 3$. (**C**) Numbers of ac^{4}C peaks in shCTR and sh*NAT10* hESCs. (**D**) Venn plot showing intersection of the ac^{4}C peaks in shCTR and sh*NAT10* hESCs. (**E**) IGV browser was used to visualize one ac^{4}C-target gene (*CCDC34*) and one nonacetylated gene (*POU5F1*). Blue and gray colors represent IP and input abundance. (**F**) Genomic distributions of the ac^{4}C peaks in shCTR and sh*NAT10* hESCs. (**G**) Distributions of the ac^{4}C peaks across transcripts. (**H**) Most significant motif in shCTR and sh*NAT10* hESCs peaks. (**I**) GO enrichment analysis for the genes with overlapped ac^{4}C peaks in shCTR and sh*NAT10* hESCs. (**J**) Comparison of percent spliced in (PSI) between shCTR and sh*NAT10* hESCs. (**K**) Comparison of percent intron retention (PIR) between shCTR and sh*NAT10* hESCs. The significance test in (J) and (K) was performed using pairwise Wilcoxon test. (**L**) Scatter density plot compares fold enrichments of ac^{4}C peaks in shCTR and sh*NAT10* hESCs. (**M**) Differential peak analysis was performed between the ac^{4}C peaks in shCTR and sh*NAT10* hESCs. ****$P < 0.0001$. 3′UTR, 3′ untranslated region; FC, fold change.

knockdown showed certain correlation, indicating that the alterations of protein levels are generally due to the changes of mRNA levels (fig. S6A). Through differential protein analysis, we identified 519 up-regulated and 541 down-regulated proteins, respectively (Fig. 5A and fig. S6B). Considering the overall reduced ac^{4}C abundance upon *NAT10* knockdown, we particularly focused on the down-regulated proteins because they could be directly regulated by ac^{4}C modification in mRNAs. Notably, analysis for these down-regulated proteins revealed significant enrichment in chromatin-associated terms, such as chromatin organization, switch/sucrose nonfermentable (SWI/SNF) superfamily complex, chromatin remodeling, and nucleosome organization (Fig. 5B and fig. S6C).

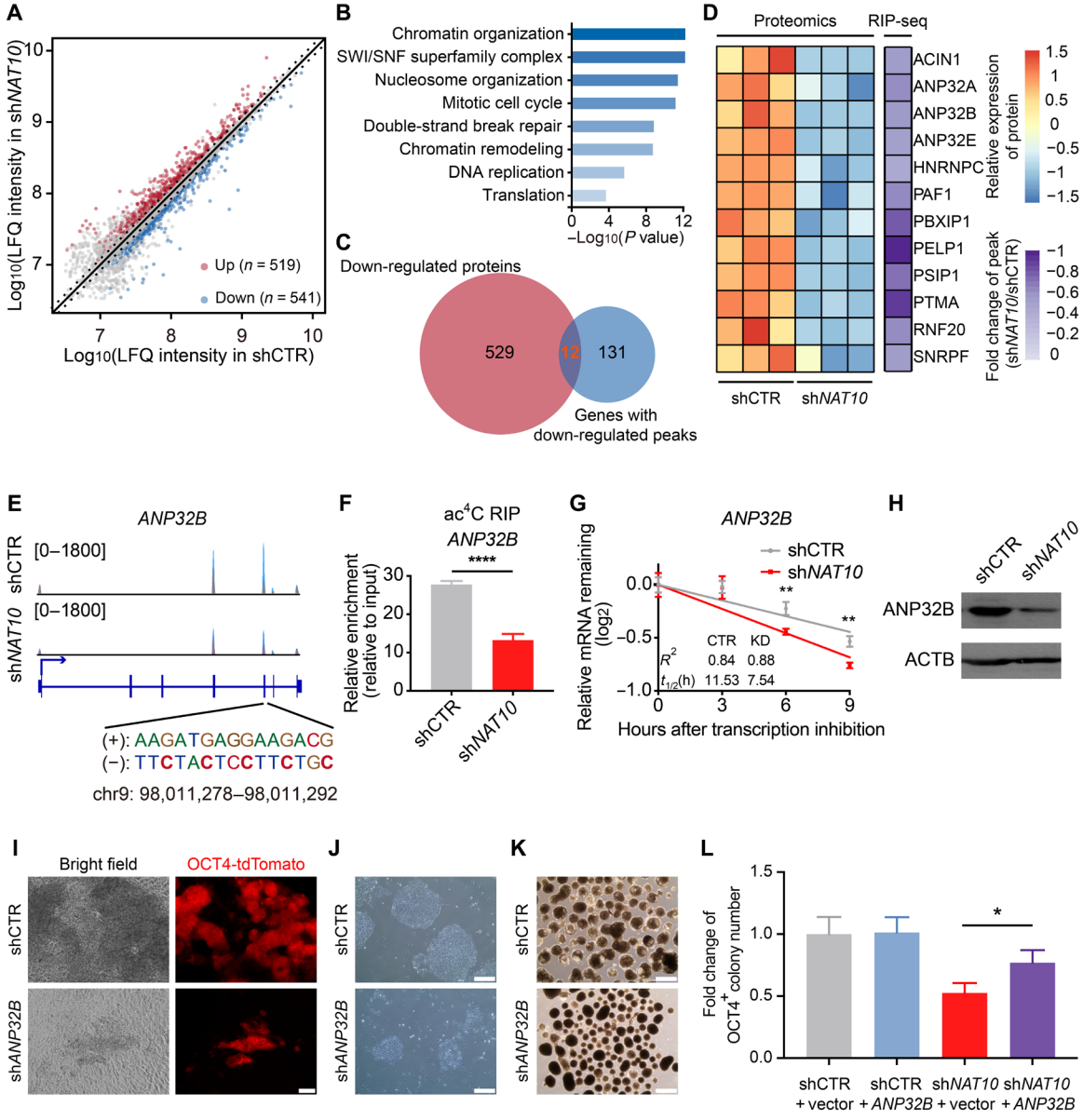

Fig. 5. Quantitative proteomics analysis identified chromatin regulators as NAT10 targets. (**A**) Dot plot displaying protein expression changes between shCTR and sh*NAT10* hESCs. (**B**) GO enrichment analysis of the down-regulated proteins. LFQ, label free quantitation. (**C**) Venn diagram showing the intersection of the down-regulated proteins and the genes with down-regulated ac^{4}C peaks. (**D**) Heatmaps displaying protein expression levels (red to blue gradient colors) of 12 overlapped genes (left) and fold changes of peaks in sh*NAT10* hESCs relative to shCTR hESCs (right). (**E**) IGV browser was used to visualize ac^{4}C peaks in chromatin regulator *ANP32B* (top) and representation of ac^{4}C motif within gene (down). (**F**) ac^{4}C-RIP-qPCR showing the ac^{4}C enrichment decreased on *ANP32B* mRNA transcript in sh*NAT10* hESCs compared with shCTR hESCs. Data are means ± SD of $N = 3$. (**G**) Decay curves for *ANP32B* in shCTR and sh*NAT10* hESCs. Data are means ± SD of $N = 3$. KD, knockdown. (**H**) Western blotting of ANP32B protein in the shCTR and sh*NAT10* hESCs. (**I**) Representation of OCT4-positive iPSC colony in shCTR cells and sh*ANP32B* cells on day 12. Scale bar, 100 μm. (**J**) Light microscopy images of day 6 colony formation for hESCs transduced with shCTR and sh*ANP32B*. Scale bars, 500 μm. (**K**) Representative images showing the morphologies of EBs from shCTR and sh*ANP32B* hESCs at day 8 of differentiation. Scale bars, 500 μm. (**L**) Pluripotent reprogramming assay of human fibroblasts after transduction with indicated lentiviruses. Data are means ± SD of $N = 3$. *$P < 0.05$, **$P < 0.01$, and ****$P < 0.0001$.

Moreover, by intersecting the down-regulated proteins and the genes with differential ac^{4}C peaks, we obtained 12 candidate genes, and, unexpectedly, nine of them are chromatin-associated, including *ANP32B*, *ANP32A*, *ANP32E*, *PAF1*, *ACIN1*, *PELP1*, *PSIP1*, *RNF20*, and *HNRNPC* (Fig. 5, C and D).

To experimentally identify the targets of NAT10, we next screened these candidate genes using functional relevance by applying the reprogramming assay. Among these nine chromatin-associated genes, *ANP32B* emerged as the strongest positive regulator for reprogramming (fig. S6D). The consistent phenotypes of NAT10 and ANP32B in reprogramming indicated ANP32B as a potential downstream target for NAT10. We therefore focused on ANP32B. We further checked the ac^{4}C peaks of *ANP32B* and confirmed the reduction of ac^{4}C peak signals after *NAT10* suppression (Fig. 5E). The ac^{4}C-RIP-qPCR result confirmed that the ac^{4}C modification level of *ANP32B* was indeed down-regulated after *NAT10* knockdown (Fig. 5F). Furthermore, NAT10-RIP-qPCR validated that NAT10 protein could directly bind with *ANP32B* transcript (fig. S6E). As the ac^{4}C modification has been reported to promote mRNA stability (*11*), we performed RNA stability assay and confirmed that stability of *ANP32B* mRNA significantly decreased after *NAT10* knockdown (Fig. 5G). Next, Western blotting showed a notable reduction of the protein level of ANP32B in sh*NAT10* hESCs (Fig. 5H). Collectively, these results molecularly connected an unexpected link between NAT10 and the ac^{4}C modification of *ANP32B*.

The function of histone chaperone ANP32B in cell fate transitions is largely unknown, although it has been linked to DNA damage response, apoptosis, and tumorigenesis (*25*). To further determine whether ANP32B contributes to the dysregulation of cell fate transitions in sh*NAT10* cells, we performed suppression and rescue experiments. First, we confirmed that suppression of *ANP32B* affected the formation of iPSC colony during pluripotent reprogramming, paralleling the effect of *NAT10* suppression (Fig. 5I). RNA-seq and RT-qPCR results showed that knockdown of *ANP32B* delayed MET and impaired the induction of pluripotent marker genes during pluripotent reprogramming (fig. S6, F and G). Particularly, the down-regulated terms included signaling pathways regulating pluripotency of stem cells and phosphatidylinositol 3-kinase–Akt signaling pathways (fig. S6H). sh*ANP32B* treatment also affected hESC proliferation (Fig. 5J and fig. S6I). Next, we checked the developmental potential of these hESCs using EB experiment, and again, sh*ANP32B* treatment impaired normal EB formation, suggesting its essential roles in differentiation (Fig. 5K). In addition, we found ectopic expression of *ANP32B* partially rescued the phenotypes of sh*NAT10* in the processes of pluripotent reprogramming, proliferation, and differentiation (Fig. 5L and fig. S6, J and K). Together, these data uncovered ANP32B as a functionally relevant target of the NAT10-ac^{4}C axis during cell fate transitions.

NAT10-ac^{4}C-ANP32B axis modulates chromatin landscape

Histone chaperone ANP32B plays critical roles in chromatin remodeling and transcription in various ways. Consistent with previous reports, our analyses demonstrated that many ANP32B's interacting proteins were chromatin regulators (e.g., CHAF1A and SMARCD1) and transcription factors (e.g., KLF5, MYC, and SOX2) (fig. S7A) (*25*–*27*). Next, we performed chromatin immunoprecipitation (ChIP)–MS, and the result also showed that the chromatin-associated binding partners of ANP32B were enriched in biological processes, including chromatin remodeling and DNA conformational change (fig. S7B). By RNA-seq and transposase-accessible chromatin using sequencing (ATAC-seq), we also found that ANP32B played important regulatory roles on gene expression and chromatin accessibility in hESCs (fig. S8, A to D). To evaluate the mechanisms of NAT10-ac^{4}C-ANP32B axis in chromatin landscape, we took integrative analyses of cleavage under targets and tagmentation (CUT&Tag), ATAC-seq, and RNA-seq data (Fig. 6A). First, we applied CUT&Tag to profile the genome-wide chromatin binding of ANP32B in shCTR and sh*NAT10* hESCs. We found that the peaks of ANP32B were significantly reduced upon *NAT10* knockdown, consistent with the reduction of ANP32B in sh*NAT10* hESCs (fig. S9, A and B). In total, 5349 down-regulated ANP32B peaks were identified and most of these down-regulated peaks (about 60%) were distributed in the promoter and gene body, suggesting that ANP32B may mainly regulate gene expression in relatively close range (fig. S9C). Second, to investigate the impact of ANP32B binding on chromatin accessibility, we performed ATAC-seq in shCTR and sh*NAT10* hESCs (fig. S9D), and after integrative analysis of ANP32B CUT&Tag and NAT10 ATAC-seq data, 961 closed loci and 1731 opened loci were identified (Fig. 6B). In addition, about 57% closed loci and 55% open loci were located in promoter regions, further indicating that in hESCs, ANP32B primarily bound proximal regulatory elements and regulated the gene expression (fig. S9E). Third, by further integrating RNA-seq data, 243 down-regulated target genes and 548 up-regulated target genes of the NAT10-ANP32B axis were identified (Fig. 6C). Notably, these down-regulated target genes were enriched in Wnt signaling pathway and mitogen-activated protein kinase signaling pathway, while up-regulated target genes were associated with p53 signaling pathway and TNF signaling pathway (Fig. 6D). We also observed that changes in chromatin accessibility and gene expression in sh*ANP32B* hESCs were correlated with *NAT10* suppression, further supporting that NAT10 and ANP32B were molecularly and functionally coupled (fig. S9, F and G).

As described above, NAT10 suppression affected chromatin states (Fig. 1, M and N), and then we applied integrative analysis. We classified four distinct clusters (C1 to C4) by ANP32B bound and chromatin status. In C1 and C3, about 80% of bound regions were in promoter regions and showed significant changes in H3K4me3 and H3K27me3 modification levels and were more correlated with gene expression (fig. S10, A to D). ANP32B partially mediated the effects of NAT10 on chromatin landscape (Fig. 6, E to H, and fig. S10, E and F). For example, NAT10-ANP32B targets with repressed chromatin status were enriched with motifs of KLF4 and LEF1 and associated with signaling by WNT and epigenetic regulation of gene expression (Fig. 6, E and F, and fig. S10E). NAT10-ANP32B targets with activated chromatin status were enriched with motifs of JUN-AP1 and AP-2γ and associated with regulation of type 1 interferon production (Fig. 6, G and H, and fig. S10F). Notably, several bivalent genes, including *SFRP1* and *NODAL* that are key regulators of Wnt and transforming growth factor–β (TGFβ) signaling pathways and important for cell fate transitions, were identified as targets of NAT10-ANP32B axis (Fig. 6, I and J, and fig. S8, A and C). In addition, we found that NAT10-ANP32B axis could orchestrate the redistribution of MLL1 and EZH2, writers for H3K4me3 and H3K27me3, and in turn regulated these histone modifications (fig. S10G). Functionally, the suppression of *SFRP1* and *NODAL* also affected pluripotent reprogramming (Fig. 6K and fig. S10H). Cumulatively, these observations concur with functions of NAT10 and ANP32B.

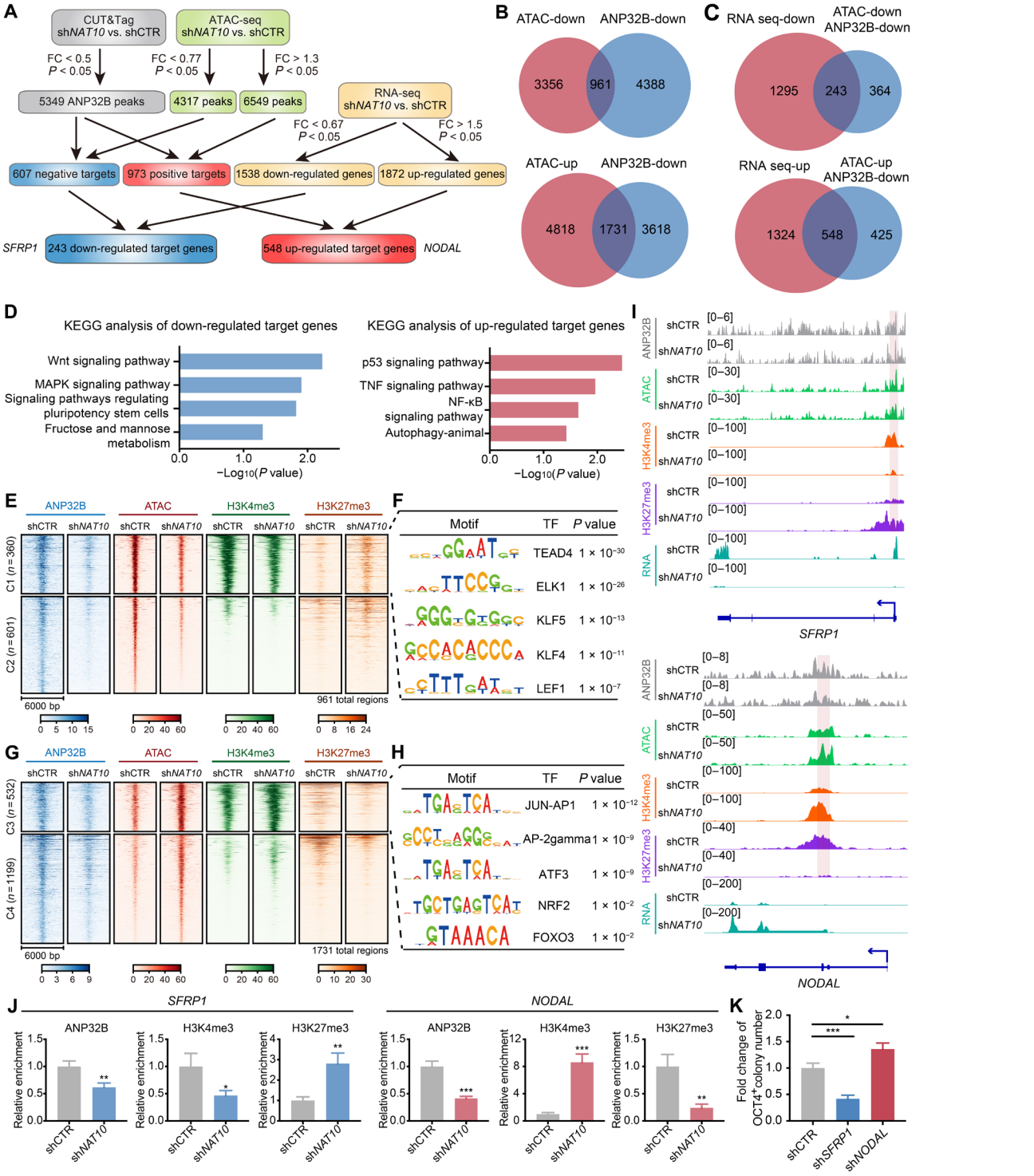

Fig. 6. NAT10-ac^{4}C-ANP32B axis modulates chromatin landscape. (**A**) Flowchart of NAT10-ac^{4}C-ANP32B downstream target gene screening. (**B**) Venn diagram showing the integrative analysis to identify potential downstream targets of NAT10-ANP32B axis. (**C**) Venn diagram showing the overlaps between differentially expressed genes and potential ANP32B targets. (**D**) KEGG pathway enrichment of NAT10-ac^{4}C-ANP32B downstream targets. (**E** and **G**) Heatmaps displaying ANP32B, ATAC, H3K4me3, and H3K27me3 signal within 6-kb window around the summit of ANP32B bound regions. (**F** and **H**) TF binding motif enriched in C1 (F) and C3 (H). (**I**) IGV browser tracks of representative NAT10-ac^{4}C-ANP32B downstream bivalent genes (*SFRP1* and *NODAL*) of showing indicated chromatin features. (**J**) CUT&Tag qPCR assay of the binding of ANP32B and the changes of H3K4me3 and H3K27me3 at *SFRP1* and *NODAL* gene loci in shCTR and sh*NAT10* hESCs. Data are means ± SD of $N = 3$. (**K**) Pluripotent reprogramming assay of human fibroblasts after transduction with shCTR, sh*SFRP1*, or sh*NODAL* lentiviruses. Data are means ± SD of $N = 3$. $^*P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

In aggregate, we found that NAT10 suppression could affect the ac^{4}C abundance of chromatin-associated genes, such as *ANP32B*, and the NAT10-ac^{4}C-ANP32B axis influenced chromatin states of their downstream target genes (e.g., *SFRP1* and *NODAL*) and gene regulatory network, broadly contributing the cell fate transitions, such as cellular reprogramming and stem cell differentiation (fig. S11).

DISCUSSION

In this study, we demonstrated that NAT10-ac^{4}C axis was essential for controlling cell fates, including stem cell differentiation and cellular reprogramming. Mechanistically, we found that NAT10-mediated ac^{4}C modifications of transcripts were involved in many important biological processes, such as chromatin remodeling, splicing, ribosome, and metabolism. We identified that epigenetic regulatory genes, including *ANP32B*, as the key downstream substrates of NAT10, and ANP32B could partially mediate the functions of NAT10 by modulating chromatin landscape in cell fate transitions.

During cell fate transitions, changes of transcriptome are highly coordinated at multiple levels. To date, our understanding of these processes mainly relies on transcription, while epigenetic and post-transcriptional regulations remain underappreciated. Besides epigenetic modulations, such as DNA and histone modifications, mRNA modifications have recently been recognized as an additional layer of controls over cellular differentiation and organ development. For example, mRNA m^{6}A dynamics during cellular reprogramming and lineage differentiation have been extensively studied in recent years (*9*, *10*). In contrast to mRNA methylation, the existence and functions of acetylation in mRNAs remain largely unknown. Here our study provides evidences for the functions of mRNA ac^{4}C in cell fate transitions.

Using tumor cell lines, a recent study identified NAT10 as the writer of ac^{4}C and described the existence of ac^{4}C modification in mammalian mRNAs. Further studies suggested that ac^{4}C modification enhances mRNA stability and translation. Here, we used hESCs, which is a very important and feasible system for studying cell fate transitions, and we found that knockdown of NAT10 affects hESC proliferation as well as spontaneous and directed lineage differentiation. In addition, NAT10 plays a crucial role in pluripotent reprogramming. Therefore, NAT10 is broadly essential for various biological processes.

Notably, we identified that epigenetic regulators, including ANP32B, as key downstream substrates of NAT10. Furthermore, we observed ac^{4}C modification in *ANP32B* transcript. In turn, ANP32B functionally mediates the effects of NAT10 during cell fate transitions. By multiple assays and integrative analyses, we found that NAT10-ac^{4}C-ANP32B axis can modulate chromatin landscape of their downstream genes, including key regulators of fate-instructive signaling pathways. Wnt and TGFβ signaling pathways are very important for cell fate determination, which further explained the broad functions of NAT10 in stem cell differentiation and somatic cell reprogramming. It should be noted that other potential molecular mechanisms of NAT10 are possible and should be explored in the future. Collectively, these studies find the essential functions of NAT10 and RNA ac^{4}C modification during cell fate transitions and will have broad implications in a variety of biological and physiological contexts.

In summary, we have uncovered a previously unrecognized role for NAT10-ac^{4}C in cell fate transitions, including reprogramming and differentiation. Mechanistically, we identified an important cross-talk between ac^{4}C-mediated epitranscriptomic regulation and ANP32B-mediated chromatin signaling. Significantly, these findings have provided a conceptual advance of cellular plasticity and can be harnessed for regenerative medicine.

MATERIALS AND METHODS

Cell culture

hESCs were maintained in hESC medium: Dulbecco's modified Eagle's medium (DMEM)/F12 (Gibco), 20% KnockOut serum replacement (KSR) (Gibco), 1× non-essential amino acids (Gibco), 1× penicillin/streptomycin (Gibco), 0.055 mM 2-mercaptoethanol (Sigma-Aldrich), and basic fibroblast growth factor (bFGF) (10 ng ml^{-1}; PeproTech). Human fibroblasts were differentiated from H9 hESC lines with the OCT4-tdTomato reporter. Human fibroblasts were cultured in DMEM, 10% fetal bovine serum (FBS) (Gibco), 1× penicillin-streptomycin, and 250 μM vitamin C (Sigma-Aldrich). Mycoplasma contamination was routinely tested.

Cell proliferation analysis

Cells were seeded at a density of 1×10^4 ml^{-1} and harvested by using 0.05% trypsin after 24, 48, 72, and 96 hours. Viable cells were counted by Cellometer (Life Technologies).

RT-qPCR assay

Total RNA was isolated using the FastPure Cell/Tissue Total RNA Isolation Kit (Vazyme) and converted to cDNA using the Reverse Transcriptase Kit (Takara). RT-qPCR was performed with the SYBR RT-PCR Kit (Takara) on CFX ConnectTM Real-Time system (Bio-Rad). Primer sequences are listed in table S2.

AP assay

After 12 days of reprogramming, cells were analyzed by AP staining using an Alkaline Phosphatase solution kit (Beyotime). Briefly, cells were fixed using 4% paraformaldehyde (PFA) for 2 min and then washed twice with phosphate-buffered saline (PBS). Subsequently, cells were incubated with the indicated bromochloroindolyl phosphate/nitro blue tetrazolium dyeing working solution for 10 min at room temperature. Cells were then washed twice with PBS to stop the reaction.

OP-Puro and EdU incorporation assays

To measure protein synthesis or DNA synthesis, hESCs were incubated for 1 hour in hESC medium supplemented with *O*-propargyl puromycin (OP-Puro) (25 mM final concentration; MedChemExpress) or 5-ethynyl-2'-deoxyuridine (EdU) (10 mM final concentration; Beyotime). After harvesting by using 0.05% trypsin, the samples were fixed for 10 min at room temperature with 4% PFA and washed with phosphate-buffered saline with Triton X-100 (PBST) three times at 1500 rpm, 5 min per time. OP-Puro or EdU labeling and staining of the hESCs were performed using the BeyoClick EdU Cell Proliferation Kit with Alexa Fluor 555 (Beyotime) according to the manufacturer's instructions.

Knockdown experiments

Lentiviral vectors (pLKO.1) carrying independent targeting or control shRNAs (sequences are shown in table S2) were transfected into HEK293T cells maintained in DMEM medium using Lipofectamine

3000 Transfection Reagent (Invitrogen). Viruses were collected and filtered 48 and 72 hours after transfection. Lentiviral infection in hESCs was performed as follows: 60 to 70% confluent hESCs were incubated with viruses and basal medium as 1:1. In addition, polybrene infection reagent (Santa Cruz Biotechnology) was added at 5 μg ml^{-1}. After 4 hours, hESC medium was used to culture cells. For cell collection, hESCs with different shRNAs were harvested by using 0.05% trypsin. Cell samples were then collected for Western blot or RT-qPCR analysis.

Human fibroblasts were infected with lentiviral particles coding for targeting or control shRNA in the presence of polybrene (5 μg ml^{-1}) in day 2 of reprogramming. sgRNA sequence was designed by CHOPCHOP, and the infection steps were the same as shRNAs. And 24 hours later, human fibroblasts were infected again.

Generation of iPSCs and reprogramming efficiency evaluation

Human fibroblasts were reprogrammed using hRep medium: basic DMEM (Gibco), 10% KSR, 10% FBS, 1× non-essential amino acids, 1× penicillin/streptomycin, 0.055 mM 2-mercaptoethanol, bFGF (10 ng ml^{-1}), 0.5 mM valproic acid (VPA), 10 μM 616452, 0.5 μM AM580, 5 μM EPZ004777, 5 μM forskolin (https://www.targetmol.cn/compound/Forskolin), 0.2 μM PD0325901, and doxycycline (https://www.targetmol.cn/compound/doxycycline) (2 μg ml^{-1}). The medium was changed every 2 days. After 12 days of reprogramming, we counted the number of OCT4-tdTomato$^+$ colony and AP^+ colony.

Lineage differentiation

For EB differentiation assay, hESCs were dissociated by scraping and blowing and then transferred to the ultralow attachment six-well plate. For definitive endoderm differentiation assay, hESCs were seeded onto six-well plates and cultured in differentiation media for 3 days and fixed for immunocytochemical analysis or harvested for RT-qPCR. The differentiation media are as follows: day 1: RPMI (Gibco), 1× penicillin-streptomycin, activin A (100 ng ml^{-1}; PeproTech), and 3 μM CHIR99021; day 2: RPMI, 1× penicillin-streptomycin, 0.2% FBS, and activin A (100 ng ml^{-1}); day 3: RPMI, 1× penicillin-streptomycin, 2% FBS, and activin A (100 ng ml^{-1}).

For pancreatic differentiation, hESCs were differentiated by a previously described protocol (*28*). The pancreatic differentiation media are as follows: days 1 to 3: The media are the same as the definitive endoderm differentiation media. Days 4 to 6: RPMI, 0.5× B27, 0.5× N2, 0.05% bovine serum albumin (BSA), 1× penicillin-streptomycin, and keratinocyte growth factor (KGF) (50 ng ml^{-1}). Days 7 to 8: DMEM, 1× B27, 0.05% BSA, 1× penicillin-streptomycin, 0.25 mM vitamin C, KGF (50 ng ml^{-1}), 0.1 μM LDN-193189, 0.1 μM GDC-0449, and 2 μM retinoic acid. Days 9 to 14: DMEM, 1× B27, 0.05% BSA, 1× penicillin-streptomycin, 0.25 mM vitamin C, 0.1 μM LDN-193189, and epidermal growth factor (50 ng ml^{-1}).

Immunostaining

Cells were fixed with 4% PFA for 10 min at room temperature and washed with PBST buffer for three times. The cells were then blocked in blocking buffer for 1 hour at room temperature and then incubated with the primary antibody overnight at 4°C. After three washes with PBST for 15 min each at room temperature, the secondary antibodies were incubated for 1 hour at room temperature. Primary and secondary antibodies are listed in table S1. In addition, cells were stained with Hoechst (1:5000) to mark nuclei.

Flow cytometry

In brief, cells were digested into single cells using 0.05% trypsin followed by washing with PBS buffer. Then, the cells were fixed by 4% PFA for 15 to 30 min and washed with PBST three times at 1500 rpm for 5 min per time. Then, cells were blocked with blocking buffer followed by incubated with primary antibodies at 4°C overnight. After washing cells using PBST for three times at 1500 rpm, cells were incubated with secondary antibodies at room temperature for 1 to 2 hours. All antibodies are detailed in table S1. FACS data were acquired by Beckman CytoFlex and analyzed by CytExpert software.

Teratoma formation

About 4×10^6 hESCs were mixed with 100 μl of Matrigel and subcutaneously injected into the dorsal flank of immunodeficient SCID Beige mice (Beijing Vital River). After 6 weeks, teratomas were surgically dissected and analyzed with hematoxylin and eosin staining.

ac^4C detection and quantification by MS

About 1 μg of total RNAs or 2 μg of poly(A) RNAs was digested by nuclease P1 (1 U) (Wako) in 30 μl of buffer containing 20 mM CH_3COONH_4 at 42°C for 2 hours. And then, 10× bacterial AP buffer and 4 U/10 μg of bacterial AP (1 U) (TOYOBO) were added for 2 hours at 37°C. Following digestion, samples were spin-filtered to remove enzymatic constituents (Millipore). The nucleosides were separated by reverse phase ultraperformance LC on a C18 column and detected by SCIEX, QTRAP 6500$^+$ LC-MS/MS triple-quadrupole LC mass spectrometer in positive electrospray ionization mode. The nucleosides were quantified by using the nucleoside-to-base ion mass transitions of 286.1 to 154.1 (ac^4C) and 244.1 to 112.0 (C). Quantification was performed by comparison with the standard curve obtained from pure nucleoside standards run on the same batch of samples. The ratio of ac^4C to C was calculated from the calibrated concentrations.

In vitro transcription of ac^4C positive and negative RNA probe

DNA oligos with T7 promoter were used as templates for in vitro transcription through PCR. The ac^4C positive and negative RNA probes were then synthesized by in vitro transcription using the HiScribe T7 Yield RNA Synthesis Kit (NEB) following protocol. Specially, when the ac^4C positive probe was synthesized, ac^4CTP instead of CTP was used. The transcribed products were digested by deoxyribonuclease I at 37°C for 1 hour to remove the template DNA and then purified with RNA clean & concentrator-25 (Zymo Research). Primers used and the sequences of the probes were listed in table S2.

Acetylated RNA immunoprecipitation and sequencing

ac^4C RNA immunoprecipitation was performed with several modifications from previous m^6A-seq protocol (*29*). In brief, 4 μg of anti-ac^4C antibody (Abcam) was prebound to 25 μl of Dynabeads protein G (Invitrogen) in PBS at room temperature for 1 hour, and then 4 μg of fragmented mRNA was spiked in with 4 pg of ac^4C-positive and ac^4C-negative probe and incubated with the antibody precoated protein G beads at 4°C for 4 hours in immunoprecipitation (IP) buffer [10 mM tris-HCl, 150 mM NaCl, and 0.1% (v/v) Igepal CA-630]. The RNA-antibody-beads complex was then washed with IP buffer for three times, ac^4C-containing RNAs were lastly eluted with ac^4CTP nucleoside. Both input and IP products were subjected to library construction using the SMARTer Stranded Total RNA-Seq Kit v2 (Takara).

RNA stability assay

shCTR and sh*NAT10* hESCs were treated with actinomycin D (Sigma-Aldrich) at a final concentration of 5 μg ml^{-1} for indicated time (0, 3, 6, and 9 hours) and collected. The total RNA was extracted by the Fast Pure Cell/Tissue Total RNA Isolation Kit and analyzed by RT-qPCR. The half-life of mRNA was calculated according to a previously reported method (*30*).

CUT&Tag assay

CUT&Tag assay was carried out similarly as described (*31*). 1×10^5 shCTR and sh*NAT10* hESCs were used for sequencing library construction. DNA was amplified using the TruePrep Index Kit V2 (Vazyme). For ANP32B library construction, the PCR cycle number was 18. For H3K4me3 and H3K27me3 library construction, the PCR cycle number was 12. The source of the reagents are listed in table S1.

RNA-seq analysis

The clean reads were mapped to human reference genome (hg38) using HISAT2 (v2.2.1) with default parameters (*32*). Samtools (v1.9) was used for sorting, indexing, and transforming reads alignment results (*33*). FeatureCounts (v2.0.1) and stringtie (v2.1.4) were exploited for calculating the raw read counts and the transcripts per million for each gene with default parameters (*34*, *35*). Differential expression analysis was performed using R package DESeq2 (*36*). Differentially expressed genes were defined on the basis of absolute expressional change higher than 1.5-fold [$|\log_2(FC)| > 0.585$] and false discovery rate (FDR) smaller than 0.05 (FDR ≤ 0.05). GO and KEGG enrichment analysis and GO network were performed using ClueGO (v2.5.8) and CluePedia (v1.5.8) with right-sided hypergeometric test (*37*, *38*).

acRIP-seq analysis

The clean reads were mapped to human reference genome (hg38) using Bowtie2 (v2.4.1) with modified parameters of "--local -k 1 --no-unal -1" (*39*). Then, sambamba (0.7.1) was used to filtered out the unmapped, duplicated, and mitochondrial reads (*40*). Last, the filtered and deduplicated bam files were exported to R package exomePeak2 with default parameters for ac^4C peak calling (*41*). To obtain high-confident ac^4C peaks, we only retained (i) peaks with at least 1.5-fold enrichment in IP relative to input and adjusted *P* value <0.05, (ii) peaks within protein-coding genes, (iii) genes with peaks should have fragments per kilobase of transcript per million fragments mapped (FPKM) higher than 1. For differential peak analysis, exomePeak2 was used to compare the relative fold enrichment of peaks in shCTR and sh*NAT10* samples, and differential ac^4C peaks were defined on the basis of at least 1.3-fold of shCTR peaks fold enrichment relative to sh*NAT10* peaks fold enrichment [$|\log_2(\text{shCTR/sh}NAT10)| > 0.378$] and adjusted *P* value < 0.05. Motif enrichment analysis was performed using Homer (v4.11) with option "-rna" (*42*). IGV and Circos were used to visualize peak abundance and chromosomal distribution, respectively (*43*, *44*).

RNA splicing analysis

The clean reads were imported to vast-tools that used VASTDB to quantify PSI and PIR in shCTR and sh*NAT10* hESCs (*45*). Only the splicing events (exon inclusion/exclusion and intron retention) with at least 20 actual supporting reads were retained for subsequent analysis.

Proteomic analysis

MS experiments were performed on a Q Exactive HF-X instrument (Thermo Fisher Scientific). Thermo Raw MS files were analyzed by MaxQuant. Uniprot was used to annotate human proteome. Label-free quantification (LFQ) intensity values were generated to quantify protein expression. Student's *t* test was used for significance test. Differential proteins were defined if *P* value < 0.05, and protein expressions have at least 1.2-fold change in shCTR and sh*NAT10* samples.

CUT&Tag analysis

The clean reads were mapped to human reference genome (hg38) using Bowtie2 (v2.4.1) with modified parameters of "--local --very-sensitive-local --no-mixed --no-discordant --phred33 -I 10 -X 700" (*39*). Macs2 (v2.1.2) with parameters "-p 1e-5 --keep-dup all –broad" was used for histone modification peak calling, and the parameters "-p 1e-5 --keep-dup all" was used for *ANP32B* peak calling (*46*). Bedtools was used for peak merging, sorting, and intersecting (*47*). To define bivalent genes, we intersected H3K4me3 and H3K27me3 peaks with minimum overlapped fraction of 10% of peak length and retained the peaks that are located at promoter regions (2-kb upstream of transcription start sites) of protein-coding genes. Bigwig files were generated with bamCoverage function in deepTools (v3.4.3) (*48*) with default parameter settings.

ATAC-seq analysis

Paired-end sequencing was conducted with the Illumina NovaSeq 6000 with read length of 150 base pairs. The clean reads were aligned to the human genome (hg38) using Bowtie2 (v2.2.5) with options: --no-discordant --no-unal --no-mixed -X 2000 --local (*39*). The sambamba (v0.7.1) software was used for removing the duplicated, unmapped, and mitochondrial reads (*40*). For ATAC-seq peak calling, parameters used were macs2 (v2.1.2) callpeak: -p 1e-5 --nomodel --shift −100 --extsize 200. The peak visualization in genome was shown by IGV software. Also, the peak annotations were performed with R package ChIPseeker (*49*), the differential ATAC peaks in shCTR and sh*NAT10* hESCs were identified by R package DESeq2 with the thresholds: absolute fold change more than 1.3 [$|\log_2(FC)| > 0.378$] and adjusted *P* value < 0.05. Bigwig files were generated with bamCoverage function in deepTools (v3.4.3) (*48*) with default parameter settings. For enhancer identification, we first selected the ATAC peaks distal to the promoters (±100 kb to ±2.5 kb of transcription start sites) and then used ROSE software with parameters -g HG38 -s 12500 -t 2500 to define the final enhancer regions.

Statistical analyses

Statistical analyses were performed by standard unpaired Student's *t* test (two-tailed, 95% confidence intervals) using GraphPad Prism 7. No statistical method was used to predetermine sample size. No samples were excluded. The experiments were not randomized. The investigators were not blinded to allocation during the experiments and outcome assessment.

Supplementary Materials

This PDF file includes:

Figs. S1 to S11

Tables S1 and S2

REFERENCES AND NOTES

1. Z. Zhu, D. Huangfu, Human pluripotent stem cells: An emerging model in developmental biology. *Development* **140**, 705–717 (2013).
2. L. Yiangou, A. D. B. Ross, K. J. Goh, L. Vallier, Human pluripotent stem cell-derived endoderm for modeling development and clinical applications. *Cell Stem Cell* **22**, 485–499 (2018).
3. T. W. Theunissen, R. Jaenisch, Molecular control of induced pluripotency. *Cell Stem Cell* **14**, 720–734 (2014).
4. R. Stadhouders, G. J. Filion, T. Graf, Transcription factors and 3D genome conformation in cell-fate decisions. *Nature* **569**, 345–354 (2019).
5. M. Frye, B. T. Harada, M. Behm, C. He, RNA modifications modulate gene expression during development. *Science* **361**, 1346–1349 (2018).
6. P. Boccaletto, B. Bagiński, MODOMICS: An operational guide to the use of the RNA modification pathways database. *Methods Mol. Biol.* **2284**, 481–505 (2021).
7. I. A. Roundtree, M. E. Evans, T. Pan, C. He, dynamic RNA modifications in gene expression regulation. *Cell* **169**, 1187–1200 (2017).
8. P. J. Batista, B. Molinie, J. Wang, K. Qu, J. Zhang, L. Li, D. M. Bouley, E. Lujan, B. Haddad, K. Daneshvar, A. C. Carter, R. A. Flynn, C. Zhou, K. S. Lim, P. Dedon, M. Wernig, A. C. Mullen, Y. Xing, C. C. Giallourakis, H. Y. Chang, m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell* **15**, 707–719 (2014).
9. T. Chen, Y. J. Hao, Y. Zhang, M. M. Li, M. Wang, W. Han, Y. Wu, Y. Lv, J. Hao, L. Wang, A. Li, Y. Yang, K. X. Jin, X. Zhao, Y. Li, X. L. Ping, W. Y. Lai, L. G. Wu, G. Jiang, H. L. Wang, L. Sang, X. J. Wang, Y. G. Yang, Q. Zhou, m(6)A RNA methylation is regulated by microRNAs and promotes reprogramming to pluripotency. *Cell Stem Cell* **16**, 289–301 (2015).
10. X. Ma, J. Cao, Z. Zhou, Y. Lu, Q. Li, Y. Jin, G. Chen, W. Wang, W. Ge, X. Chen, Z. Hu, X. Shu, Q. Deng, J. Pu, C. Liang, J. Fu, J. Liu, S. Zhu, N(6)-methyladenosine modification-mediated mRNA metabolism is essential for human pancreatic lineage specification and islet organogenesis. *Nat. Commun.* **13**, 4148 (2022).
11. D. Arango, D. Sturgill, N. Alhusaini, A. A. Dillman, T. J. Sweet, G. Hanson, M. Hosogane, W. R. Sinclair, K. K. Nanan, M. D. Mandler, S. D. Fox, T. T. Zengeya, T. Andresson, J. L. Meier, J. Coller, S. Oberdoerffer, acetylation of cytidine in mRNA promotes translation efficiency. *Cell* **175**, 1872–1886.e24 (2018).
12. D. Arango, D. Sturgill, R. Yang, T. Kanai, P. Bauer, J. Roy, Z. Wang, M. Hosogane, S. Schiffers, S. Oberdoerffer, Direct epitranscriptomic regulation of mammalian translation initiation through N4-acetylcytidine. *Mol. Cell* **82**, 2797–2814.e11 (2022).
13. K. Tsai, A. A. Jaguva Vasudevan, C. Martinez Campos, A. Emery, R. Swanstrom, B. R. Cullen, Acetylation of cytidine residues boosts HIV-1 gene expression by increasing viral RNA stability. *Cell Host Microbe* **28**, 306–312.e6 (2020).
14. Y. Zhang, Y. Jing, Y. Wang, J. Tang, X. Zhu, W. L. Jin, Y. Wang, W. Yuan, X. Li, X. Li, NAT10 promotes gastric cancer metastasis via N4-acetylated COL5A1. *Signal Transduct. Target. Ther.* **6**, 173 (2021).
15. G. Wang, M. Zhang, Y. Zhang, Y. Xie, J. Zou, J. Zhong, Z. Zheng, X. Zhou, Y. Zheng, B. Chen, C. Liu, NAT10-mediated mRNA N4-acetylcytidine modification promotes bladder cancer progression. *Clin. Transl. Med.* **12**, e738 (2022).
16. T. I. Oh, Y. M. Lee, B. O. Lim, J. H. Lim, Inhibition of NAT10 suppresses melanogenesis and melanoma growth by attenuating microphthalmia-associated transcription factor (MITF) expression. *Int. J. Mol. Sci.* **18**, (2017).
17. Y. Tan, J. Zheng, X. Liu, M. Lu, C. Zhang, B. Xing, X. Du, Loss of nucleolar localization of NAT10 promotes cell migration and invasion in hepatocellular carcinoma. *Biochem. Biophys. Res. Commun.* **499**, 1032–1038 (2018).
18. Z. Liu, X. Liu, Y. Li, P. Ren, C. Zhang, L. Wang, X. Du, B. Xing, miR-6716-5p promotes metastasis of colorectal cancer through downregulating NAT10 expression. *Cancer Manag. Res.* **11**, 5317–5332 (2019).
19. D. Larrieu, S. Britton, M. Demir, R. Rodriguez, S. P. Jackson, Chemical inhibition of NAT10 corrects defects of laminopathic cells. *Science* **344**, 527–532 (2014).
20. D. Larrieu, E. Viré, S. Robson, S. Y. Breusegem, T. Kouzarides, S. P. Jackson, Inhibition of the acetyltransferase NAT10 normalizes progeric and aging cells by rebalancing the Transportin-1 nuclear import pathway. *Sci. Signal.* **11**, (2018).
21. G. Balmus, D. Larrieu, A. C. Barros, C. Collins, M. Abrudan, M. Demir, N. J. Geisler, C. J. Lelliott, J. K. White, N. A. Karp, J. Atkinson, A. Kirton, M. Jacobsen, D. Clift, R. Rodriguez, D. J. Adams, S. P. Jackson, Targeting of NAT10 enhances healthspan in a mouse model of human accelerated aging syndrome. *Nat. Commun.* **9**, 1700 (2018).
22. B. E. Bernstein, T. S. Mikkelsen, X. Xie, M. Kamal, D. J. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath, R. Jaenisch, A. Wagschal, R. Feil, S. L. Schreiber, E. S. Lander, A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
23. A. Sas-Chen, J. M. Thomas, D. Matzov, M. Taoka, K. D. Nance, R. Nir, K. M. Bryson, R. Shachar, G. L. S. Liman, B. W. Burkhart, S. T. Gamage, Y. Nobe, C. A. Briney, M. J. Levy, R. T. Fuchs, G. B. Robb, J. Hartmann, S. Sharma, Q. Lin, L. Florens, M. P. Washburn, T. Isobe, T. J. Santangelo, M. Shalev-Benami, J. L. Meier, S. Schwartz, Dynamic RNA acetylation revealed by quantitative cross-evolutionary mapping. *Nature* **583**, 638–643 (2020).
24. G. Jin, M. Xu, M. Zou, S. Duan, The processing, gene regulation, biological functions, and clinical relevance of N4-Acetylcytidine on RNA: A systematic review. *Mol. Ther. Nucleic Acids* **20**, 13–24 (2020).
25. M. Yu, Y. Qu, H. Zhang, X. Wang, Roles of ANP32 proteins in cell biology and viral replication. *Animal Diseases* **2**, 22 (2022).
26. Y. Munemasa, T. Suzuki, K. Aizawa, S. Miyamoto, Y. Imai, T. Matsumura, M. Horikoshi, R. Nagai, Promoter region-specific histone incorporation by the novel histone chaperone ANP32B and DNA-binding factor KLF5. *Mol. Cell. Biol.* **28**, 1171–1181 (2008).
27. N. Tochio, T. Umehara, Y. Munemasa, T. Suzuki, S. Sato, K. Tsuda, S. Koshiba, T. Kigawa, R. Nagai, S. Yokoyama, Solution structure of histone chaperone ANP32B: Interaction with core histones H3-H4 through its acidic concave domain. *J. Mol. Biol.* **401**, 97–114 (2010).
28. A. Rezania, J. E. Bruin, P. Arora, A. Rubin, I. Batushansky, A. Asadi, S. O'Dwyer, N. Quiskamp, M. Mojibian, T. Albrecht, Y. H. Yang, J. D. Johnson, T. J. Kieffer, Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nat. Biotechnol.* **32**, 1121–1133 (2014).
29. D. Dominissini, S. Moshitch-Moshkovitz, M. Salmon-Divon, N. Amariglio, G. Rechavi, Transcriptome-wide mapping of N6-methyladenosine by m6A-seq based on immunocapturing and massively parallel sequencing. *Nat. Protoc.* **8**, 176–189 (2013).
30. C. Y. Chen, N. Ezzeddine, A. B. Shyu, Messenger RNA half-life measurements in mammalian cells. *Methods Enzymol.* **448**, 335–357 (2008).
31. H. S. Kaya-Okur, S. J. Wu, C. A. Codomo, E. S. Pledger, T. D. Bryson, J. G. Henikoff, K. Ahmad, S. Henikoff, CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* **10**, 1930 (2019).
32. D. Kim, J. M. Paggi, C. Park, C. Bennett, S. L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
33. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, S., The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
34. Y. Liao, G. K. Smyth, W. Shi, featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
35. M. Pertea, G. M. Pertea, C. M. Antonescu, T. C. Chang, J. T. Mendell, S. L. Salzberg, StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
36. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
37. G. Bindea, B. Mlecnik, H. Hackl, P. Charoentong, M. Tosolini, A. Kirilovsky, W. H. Fridman, F. Pages, Z. Trajanoski, J. Galon, ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093 (2009).
38. G. Bindea, J. Galon, B. Mlecnik, CluePedia Cytoscape plugin: Pathway insights using integrated experimental and in silico data. *Bioinformatics* **29**, 661–663 (2013).
39. E. Giannoulatou, S. H. Park, D. T. Humphreys, J. W. Ho, Verification and validation of bioinformatics software without a gold standard: A case study of BWA and Bowtie. *BMC Bioinformatics* **15**, S15 (2014).
40. A. Tarasov, A. J. Vilella, E. Cuppen, I. J. Nijman, P. Prins, Sambamba: Fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).
41. J. Meng, X. Cui, M. K. Rao, Y. Chen, Y. Huang, Exome-based analysis for RNA epigenome sequencing data. *Bioinformatics* **29**, 1565–1567 (2013).
42. S. Heinz, C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, J. X. Cheng, C. Murre, H. Singh, C. K. Glass, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
43. J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, J. P. Mesirov, Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
44. M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, M. A. Marra, Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).
45. J. Tapial, K. C. H. Ha, T. Sterne-Weiler, A. Gohr, U. Braunschweig, A. Hermoso-Pulido, M. Quesnel-Vallières, J. Permanyer, R. Sodaei, Y. Marquez, L. Cozzuto, X. Wang, M. Gómez-Velázquez, T. Rayon, M. Manzanares, J. Ponomarenko, B. J. Blencowe, M. Irimia, An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. *Genome Res.* **27**, 1759–1768 (2017).
46. Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, X. S. Liu, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
47. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
48. F. Ramírez, F. Dündar, S. Diehl, B. A. Grüning, T. Manke, deepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
49. G. Yu, L. G. Wang, Q. Y. He, ChIPseeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).

Acknowledgments: We thank D. Fang and Y. Zhang for the help in the ATAC-seq experiment. We thank C. Ye, X. Shu, and X. Jiang for the assistance in the LC-MS/MS experiment. We thank X.

Zhang for the assistance in the IP-MS experiment. **Funding:** This work was supported by the grants from National Natural Science Foundation of China (nos. 32270781, 31970818, 22022702, and 21977087), the National Key Research and Development Program of China (nos. 2016YFC1305300, 2022YFA1103702, and 2017YFA0506800), the Outstanding Youth Fund of Zhejiang Province (no. R17C120002), Zhejiang Provincial Natural Science Foundation of China (no. LZ23B020004), and MOE Key Laboratory of Macromolecular Synthesis and Functionalization, Zhejiang University (no. 2022MSF04). **Author contributions:** Conceptualization: S.Z., J.L., Z.H., and J.C. Methodology: S.Z., J.L., Z.H., Y. Lu, J.C., L.L., X.C., Z.Z., J.P., G.C., X.M., Q.D., Y.J., L.J., Y. Li, and T.L. Data curation: Z.H., Y. Lu, J.C., S.Z., J.L., and L.L.. Investigation: Z.H., Y. Lu, J.C., S.Z., J.L., and L.L.. Visualization: S.Z., J.L., Z.H., and J.C. Supervision: S.Z. and J.L. Validation: Z.H. and J.C. Writing—original draft: S.Z., J.L., Z.H., L.L., and J.C. Writing—review and editing: S.Z., J.L., Z.H., Y. Lu, and J.C. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Sequencing data have been deposited in the NCBI Gene Expression Omnibus (GEO) database with the GEO accession number GSE210504. The accession number for the MS proteomics data reported here is ProteomeXchange Consortium via the PRIDE partner repository: PXD036888. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 29 March 2023
Accepted 14 December 2023
Published 12 January 2024
10.1126/sciadv.adh9871