Article

The RNA m⁶A reader YTHDC1 silences retrotransposons and guards ES cell identity

Check for updates

Jiadong Liu[1,2,3,11], Mingwei Gao[1,2,3,11], Jiangping He[4,11], Kaixin Wu[1,5,6], Siyuan Lin[1,2], Lingmei Jin[1,2,3], Yaping Chen[1,2,3], He Liu[4,7], Junjie Shi[1,2,3], Xiwei Wang[4], Lei Chang[4], Yingying Lin[8], Yu-Li Zhao[9], Xiaofei Zhang[1,2,4], Man Zhang[4], Guan-Zheng Luo[9], Guangming Wu[4], Duanqing Pei[1,4,10], Jie Wang[1,2], Xichen Bao[1,2,4] & Jiekai Chen[1,2,3,4,7 ✉]

The RNA modification N^6-methyladenosine (m⁶A) has critical roles in many biological processes[1,2]. However, the function of m⁶A in the early phase of mammalian development remains poorly understood. Here we show that the m⁶A reader YT521-B homology-domain-containing protein 1 (YTHDC1) is required for the maintenance of mouse embryonic stem (ES) cells in an m⁶A-dependent manner, and that its deletion initiates cellular reprogramming to a 2C-like state. Mechanistically, YTHDC1 binds to the transcripts of retrotransposons (such as intracisternal A particles, ERVK and LINE1) in mouse ES cells and its depletion results in the reactivation of these silenced retrotransposons, accompanied by a global decrease in SETDB1-mediated trimethylation at lysine 9 of histone H3 (H3K9me3). We further demonstrate that YTHDC1 and its target m⁶A RNAs act upstream of SETDB1 to repress retrotransposons and *Dux*, the master inducer of the two-cell stage (2C)-like program. This study reveals an essential role for m⁶A RNA and YTHDC1 in chromatin modification and retrotransposon repression.

m⁶A is the most abundant modification in messenger RNAs (mRNAs) and non-coding RNAs (ncRNAs) and is recognized by readers (m⁶A-binding proteins) to regulate biological processes[1,2]. Five YT521-B homology (YTH) domain family proteins and several additional proteins have been identified to selectively bind to RNAs and impact their metabolism in an m⁶A-dependent manner[2–4]. Specifically, biochemical investigations revealed structural insights into YTH-domain mediated direct m⁶A recognition[5,6]. Notably, the nuclear reader YTHDC1 exhibits a special and critical function, as its depletion leads to early embryonic lethality in mice, similarly to *Mettl3* deficiency[7,8]. By contrast, mice deficient in other YTH readers[9–12] or known YTHDC1-interacting genes such as *Zcchc8* do not exhibit early embryonic lethality[13,14]. These results suggest that YTHDC1 is responsible for unidentified m⁶A-dependent mechanisms besides post-transcriptional regulation during embryonic development. However, how YTHDC1 regulates chromatin events and consequently contributes to embryonic developmental transition remains unclear.

YTHDC1 maintains mouse ES cell identity

To investigate the unique role of YTHDC1, we constructed *Ythdc1* conditional knockout (cKO) mouse ES cells by inserting loxP sequences around *Ythdc1* exons 7–9, spanning the m⁶A-binding site of the YTH domain (Fig. 1a). After 4 days of treatment by 4-hydroxytamoxifen (4OHT), CreERT-expressing cKO cells were validated for *Ythdc1* knockout (KO) by PCR-based genotyping and western blotting (Fig. 1b, Extended Data Fig. 1a). Notably, *Ythdc1*-KO cells showed severe defects in proliferation and increasing apoptosis after being passaged (Extended Data Fig. 1b–d), demonstrating that *Ythdc1* is required for mouse ES cell maintenance. Then, we performed transcriptomic analysis of the *Ythdc1*-KO mouse ES cells and identified more than 3,000 significantly altered genes and transposable elements (TEs) (Fig. 1c). Of note, 2C-like cell markers[15–17], the *Zscan4* family, *Dux* and *MERVL* were among the genes upregulated upon *Ythdc1* KO (Fig. 1c). Consistently, by comparing the RNA-sequencing (RNA-seq) data with pre-defined 2C and early embryonic developmental gene sets, we observed enrichment of 2C genes upon *Ythdc1* KO (Fig. 1d, Extended Data Fig. 1e, f), suggesting that *Ythdc1* depletion induces 2C-like state transition at the transcriptional level. Analysis by quantitative PCR (qPCR) with reverse transcription (RT–qPCR) further validated the activation of 2C-specific genes and TEs (Fig. 1e, Extended Data Fig. 1g). Additionally, *Ythdc1* KO leads to the suppression of a pluripotency marker, NANOG, featuring 2C-like state transition[15] (Extended Data Fig. 1h). To further verify the 2C-like state induction, we generated two Rosa26-targeted MERVL–tdTomato[18] knock-in *Ythdc1*-cKO mouse ES cell lines; around 30% of these cells were labelled with tdTomato upon 4OHT treatment (Fig. 1f, g, Extended Data Fig. 1i). YTHDC1 is known to specifically recognize m⁶A-modified RNAs through the YTH domain[5,6]. We

[1]CAS Key Laboratory of Regenerative Biology, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, China. [2]Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, Guangzhou, China. [3]University of Chinese Academy of Sciences, Beijing, China. [4]Center for Cell Lineage and Atlas (CCLA), Bioland Laboratory, Guangzhou Regenerative Medicine and Health GuangDong Laboratory, Guangzhou, China. [5]Central Laboratory, The Sixth Affiliated Hospital of Guangzhou Medical University, Qingyuan People's Hospital, Qingyuan, China. [6]Department of Developmental Biology, School of Basic Medical Sciences, Southern Medical University, Guangzhou, China. [7]Joint School of Life Sciences, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou Medical University, Guangzhou, China. [8]School of Medicine, Sun Yat-Sen University, Shenzhen, China. [9]MOE Key Laboratory of Gene Function and Regulation, School of Life Sciences, Sun Yat-Sen University, Guangzhou, China. [10]School of Life Sciences, Westlake University, Hangzhou, China. [11]These authors contributed equally: Jiadong Liu, Mingwei Gao, Jiangping He. ✉e-mail: chen_jiekai@gibh.ac.cn
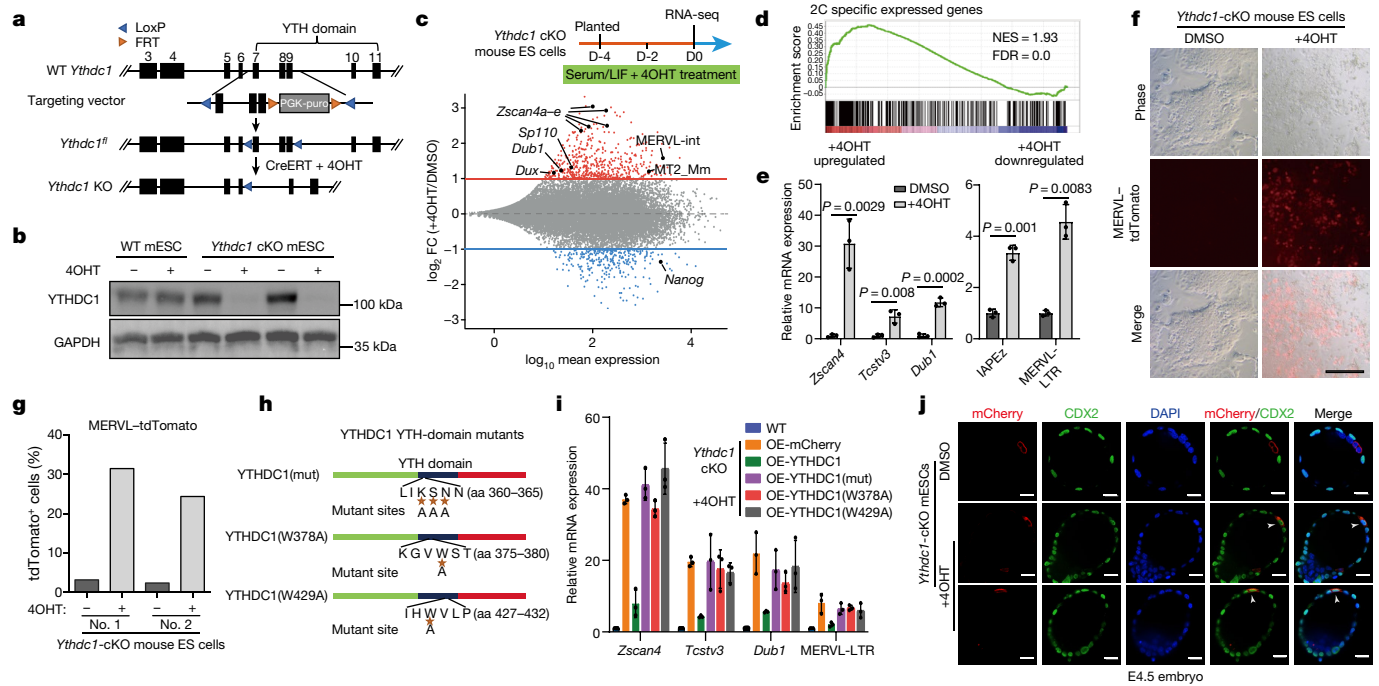
Fig. 1 | YTHDC1 depletion induces 2C-like state transition. a, Schematic showing the *Ythdc1* cKO strategy using the Cre-loxP system in mouse ES cells. KO is initiated upon 4OHT addition. **b**, Western blot analysis showing expression of YTHDC1 in WT and *Ythdc1*-cKO mouse ES cells after 4OHT treatment. Immunoblotting was repeated at least three times with similar results. **c**, Top, schematic depicting the timeline of DMSO or 4OHT treatment for RNA-seq analysis. Bottom, MA plot showing changes in expression of genes and TEs (by subfamily) upon *Ythdc1* KO. Horizontal red and blue lines define log2 fold change (FC) >1 and <−1, respectively. **d**, Gene set enrichment analysis (GSEA) illustrates the upregulation of 2C genes upon loss of *Ythdc1* (+4OHT). **e**, RT–qPCR data demonstrate the upregulation of 2C genes and TEs following *Ythdc1*-KO. Data are mean ± s.d. of three independent experiments. *P* values determined by two-sided Student's *t*-test. **f**, Phase-contrast and fluorescence images display the expression of MERVL-tdTomato reporter upon *Ythdc1* KO. Images are representative of at least two independent experiments. Scale bars, 250 μm. **g**, A column graph quantifies the percentages of MERVL–tdTomato+ cells upon *Ythdc1* KO (*n* = 2 independent experiments). **h**, Schematic showing the location of m6A-binding-site mutations in the YTH domain of YTHDC1. aa, amino acids. **i**, RT–qPCR detection of 2C genes and TEs in the rescue assays. Data are mean ± s.d. of three independent experiments. **j**, E4.5 blastocysts developing from morula injected with *Ythdc1* cKO mouse ES cells treated with DMSO or 4OHT were stained for mCherry and CDX2. Images are representative of at least two independent experiments. Scar bars, 20 μm.

constructed three known YTH m6A-binding-site mutants[5–7] of YTHDC1 (K362A/S363A/N364A, W378A and W429A) (Fig. 1h). We overexpressed wild-type (WT) or YTH m6A-binding mutants of *Ythdc1* and found that only WT *Ythdc1* rescues the activation of the 2C-like program (Fig. 1i, Extended Data Fig. 1j), indicating that the 2C-like transition is initiated in an m6A-dependent manner.

To examine whether the 2C-like state transition contributes to the reduced viability of *Ythdc1*-KO, we stained these 2C-reporter cells with EdU and DAPI. Both naturally occurring and *Ythdc1*-KO-induced 2C-like cells exhibited significantly reduced EdU staining and the disappearance of the chromatin-condensed mitotic phase (Extended Data Fig. 1k–m), indicating that the 2C-like state impairs cell proliferation. However, only about 30% of *Ythdc1*-KO cells were positive for the 2C reporter (Fig. 1g), but the 2C-negative *Ythdc1*-KO cells also showed a decreased EdU ratio (Extended Data Fig. 1l). In addition, overexpression of WT *Ythdc1*, but not the YTH-domain-mutant forms, rescued the proliferation deficiency of *Ythdc1*-KO mouse ES cells (Extended Data Fig. 1n, o). These data suggest that the vital role of YTHDC1 in mouse ES cell maintenance is dependent on m6A binding, through both 2C-like transition-dependent and -independent pathways.

We then performed a chimera assay to examine the developmental potential of *Ythdc1*-KO-induced 2C-like cells. *Ythdc1*-KO cells are able to incorporate into trophectoderm with expression of the trophectoderm marker CDX2 (Fig. 1j, Extended Data Fig. 2a–c), demonstrating the acquisition of expanded developmental potential. However, we did not detect surviving *Ythdc1*-KO cells in E6.5 injected embryos (data not shown). Given that *Ythdc1* depletion leads to either early embryonic lethality[7] or mouse ES cell proliferation defects, we proposed that the

Ythdc1-KO-induced 2C-like cells are not as functionally competent as those that emerge naturally from WT mouse ES cells. We also used reversible short interfering RNA (siRNA) to knock down (KD) *Ythdc1* (Extended Data Fig. 2d, e), and showed that *Ythdc1*-KD cells confer both embryonic and extraembryonic (ExEm) chimeric competency in E6.5 embryos (Extended Data Fig. 2f, g). Together, these observations demonstrate that *Ythdc1* relies on its m6A-binding activity to maintain ES cell identity and restrict the transition back to the 2C-like state.

YTHDC1–SETDB1 represses retrotransposons

To identify target RNAs of YTHDC1, we knocked in a Halo tag at the N terminus of the endogenous *Ythdc1* gene to enhance the specificity of YTHDC1 RNA immunoprecipitation sequencing (RIP-seq). We identified approximately 20,000 peaks (Fig. 2a, Extended Data Fig. 3a). We also performed m6A RIP-seq, which enabled us to compare the results with four previously published datasets[8,14,19,20] in relation to known m6A features such as mRNA distribution and consensus motifs (Extended Data Fig. 3b–d). Specifically, we observed m6A peaks deposited on non-coding RNAs and TE-derived RNAs such as L1Md_T (Extended Data Fig. 3e, f). The YTHDC1 RIP-seq peaks were enriched for the DRACH motif and *Neat1*, a known YTHDC1 and m6A target locus[21] (Extended Data Fig. 3g, h). Compared with the input lysate, YTHDC1 RIP-seq peaks were preferentially located at TE loci (Extended Data Fig. 3i). Among five independent m6A RIP-seq datasets, 45 TE families were consistently marked with convincing m6A signal (Extended Data Fig. 3j). Indeed, YTHDC1 bound to the majority of these m6A-marked TE families (36 out of 45; Extended Data Fig. 3k). Retrotransposons, representing the vast
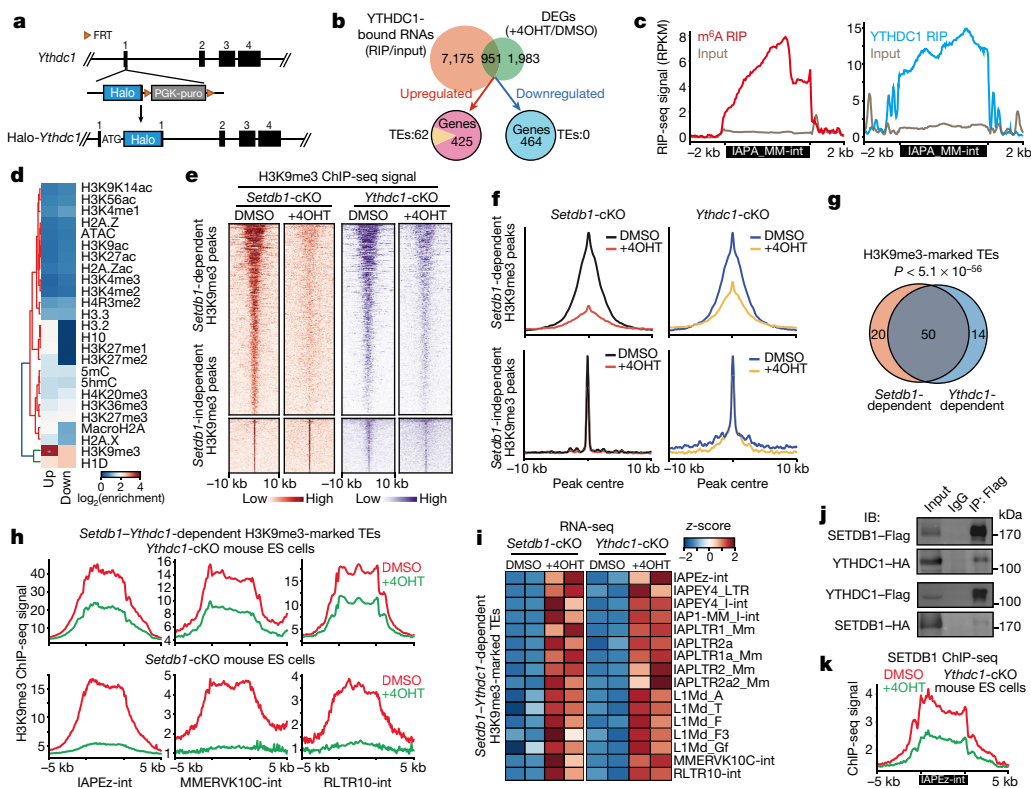
Fig. 2 | YTHDC1 mediates *Setdb1* H3K9me3-dependent retrotransposon repression. a, Schematic showing the strategy of Halo-tag knockin to the *Ythdc1* genomic locus in mouse ES cells. **b**, Top, Venn diagram showing overlapping transcripts between YTHDC1-bound RNAs from RIP-seq data (*n* = 2) and differentially-expressed RNAs from *Ythdc1*-cKO RNA-seq. Bottom, pie charts showing upregulated and downregulated common transcripts from genes and TEs. RIP, RNA immunoprecipitation. **c**, Read count tag density pileups of YTHDC1–m⁶A RIP (*n* = 2) profile for indicated TEs. RPKM, reads per kilobase of transcript, per million mapped reads. **d**, Heat map showing the enrichment of chromatin marks and DNA methylation in the up- and downregulated gene and TE clusters. The fold enrichment was calculated by the ratio of the observed over expected counts of the modified transcripts among randomly selected genomic genes and TEs for each epigenetic mark. A threshold of eightfold was applied to define an enriched chromatin mark (indicated by an asterisk). **e**, Heat

maps illustrating the density of H3K9me3 ChIP-seq reads in *Setdb1* and *Ythdc1*-cKO mouse ES cells. The H3K9me3 ChIP-seq peaks were divided into *Setdb1*-dependent and *Setdb1*-independent peaks on the basis of the difference in *Setdb1*-KO versus control. **f**, Read-count tag density pileups of H3K9me3 profiles on *Setdb1*-dependent and -independent peaks in **e**. **g**, Venn diagram showing the overlap between *Setdb1*- and *Ythdc1*-dependent TEs modified by H3K9me3. One-tailed Fisher's exact test. **h**, Read-count tag density pileups of H3K9me3 profile (from two replicates) on indicated TEs upon loss of *Ythdc1* or *Setdb1*. **i**, Heat map showing the expression of indicated TEs upon loss of *Ythdc1* or *Setdb1*. **j**, Co-immunoprecipitation assays demonstrating that YTHDC1 interacts with SETDB1. IB, immunoblot; IP, immunoprecipitation. The experiment was repeated at least three times with similar results. HA, haemagglutinin tag. **k**, Read-count tag density pileups of SETDB1 binding for indicated TEs upon *Ythdc1* depletion.

majority of TE-derived sequences, encompass endogenous retroviral elements and LINE1. These two families are activated specifically at the 2C stage, but must be repressed for the transition into blastocysts and later stages of embryonic development[15,22,23]. Indeed, we show that among the YTHDC1-targeted transcripts with differential expression upon *Ythdc1*-KO, all 62 differentially expressed TE-family-derived transcripts are upregulated, including m⁶A modified IAPA_MM-int from intracisternal A-type particle (IAP) and L1Md_T from LINE1 (Fig. 2b, c, Extended Data Fig. 3l). The genomic locations of these upregulated transcripts showed significant enrichment of H3K9me3 (Fig. 2d), suggesting that YTHDC1 represses retrotransposons via H3K9me3. These YTHDC1-mediated biological processes resemble the function of SETDB1, a H3K9me3 methyltransferase that hinders TE expression and 2C-like transition[18,24].

Genome-wide H3K9me3 chromatin immunoprecipitation followed by sequencing (ChIP–seq) analysis further supports the function of YTHDC1 in regulation of H3K9me3. Specifically, *Ythdc1* depletion reduces SETDB1-mediated H3K9me3 but has no effect on *Setdb1*-independent H3K9me3 (Fig. 2e, f, Extended Data Fig. 3m). As previously reported, most *Setdb1*-dependent H3K9me3 marks are deposited on TE regions[18]. Of note, the YTHDC1-dependent H3K9me3-marked TE targets largely overlapped with those regulated by SETDB1 (Fig. 2g, h, Extended Data

Fig. 3n). Moreover, depletion of either *Ythdc1* or *Setdb1* enhanced the global expression of these *Setdb1* and *Ythdc1* co-dependent H3K9me3 modified TEs (Fig. 2i). As YTHDC1 is known to regulate RNA decay[14], 4sUDRB-seq further delineated that *Ythdc1* KO enhanced nascent RNA transcription of the targeted TEs and genes (Extended Data Fig. 3o). *Ythdc1* depletion did not influence protein levels of SETDB1 (Extended Data Fig. 3p), and co-immunoprecipitation revealed physical interactions between YTHDC1 and SETDB1 (Fig. 2j), suggesting that m⁶A-targeted YTHDC1 recruits SETDB1 to methylate H3K9 at designated TEs. For example, *Ythdc1* KO attenuates SETDB1 occupation on IAPEz-int, a SETDB1–YTHDC1-dependent H3K9me3-marked TE (Fig. 2k, Extended Data Fig. 3q). Together, this biochemical, epigenomic and functional evidence strongly supports the co-regulation of TE silencing by YTHDC1 and SETDB1 via H3K9me3.

YTHDC1 targets chromatin via m⁶A TE RNA

As the deposition of H3K9me3 on retrotransposons requires chromatin targeting, we further investigated how YTHDC1 associates with chromatin and mediates SETDB1-dependent H3K9me3. We performed YTHDC1 ChIP-seq and identified 18,525 YTHDC1-binding peaks, of which 12,933 were also marked by H3K9me3 (Fig. 3a, Extended Data Fig. 4a, b). These
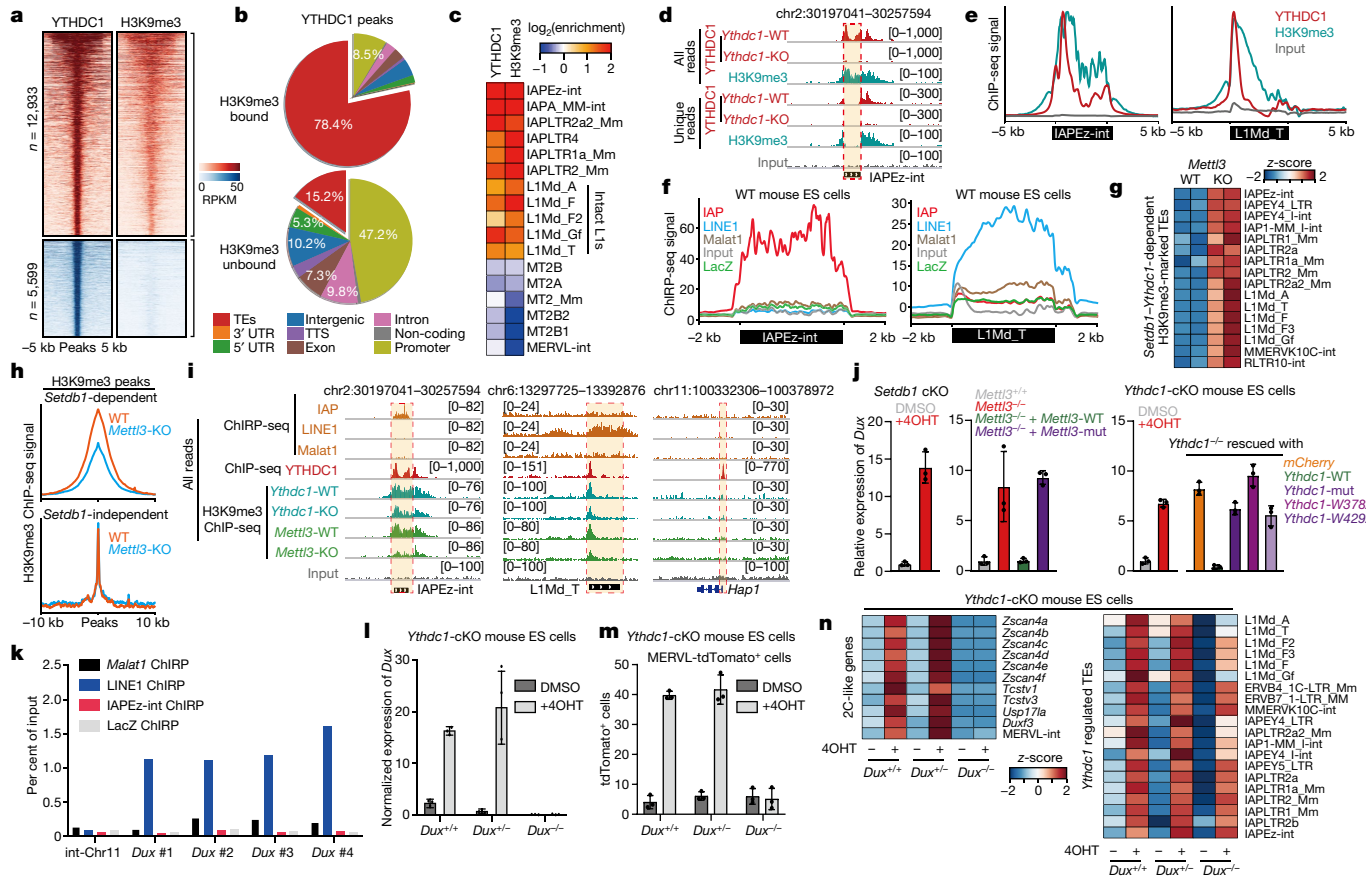
Fig. 3 | m⁶A-YTHDC1 regulates retrotransposon repression and
***Dux*-dependent 2C-like transition. a**, Heat maps showing the density of
YTHDC1 and H3K9me3 ChIP-seq reads on YTHDC1-binding peaks. **b**, Pie charts
showing the genomic distribution of the peaks identified by m⁶A and YTHDC1
ChIP-seq. **c**, Heat map showing co-enrichment of YTHDC1 and H3K9me3 on TE
families. **d**, Genome browser example of a IAPEz-int element for indicated
ChIP-seq data. All reads, using all mapped reads; unique reads, using only the
unique mapped reads. **e**, Read count tag density pileups of YTHDC1 and
H3K9me3 for IAPEz-int and the intact L1Md_T elements. **f**, Read-count tag
density pileups of IAP–LINE1 ChIRP for indicated TEs ($n = 2$). **g**, Heat map of
RNA-seq data showing the expression changes of indicated TEs upon *Mettl3*
depletion ($n = 2$). **h**, Read-count tag density pileups of H3K9me3 profiles upon
loss of *Mettl3* ($n = 2$) on *Setdb1*-dependent and independent peaks from Fig. 2e.
i, Selected genomic views of ChIRP and ChIP-seq data for the indicated TEs and
genes. **j**, Expression of *Dux* upon depletion of *Setdb1*, *Mettl3* or *Ythdc1* and
rescue, measured by qPCR using primer sets *Dux* #1–4. Data are mean ± s.d. of
three independent experiments. **k**, ChIRP–qPCR enrichment in WT mouse ES
cells at the indicated DNA loci using biotinylated probes against *Malat1*, LINE1
and IAPEz-int RNA or LacZ control ($n = 2$ independent experiments). **l**, Bar chart
showing the expression of *Dux* upon *Dux* KO in *Ythdc1* cKO mouse ES cell. Data
are mean ± s.d. from $n = 3$ independent experiments. **m**, Percentages of
MERVL-tdTomato-positive cells upon *Dux* KO in *Ythdc1*-cKO mouse ES cells
treated with DMSO or 4OHT. Data are mean ± s.d. from $n = 3$ independent
experiments. **n**, Heat map showing expression changes of the indicated genes
and TEs upon *Dux* depletion in *Ythdc1*-cKO mouse ES cells treated with DMSO or
4OHT.

YTHDC1–H3K9me3 co-bound peaks mainly localize to TE regions, such
as IAPEz-int and L1Md_T, whereas the YTHDC1-binding peaks not bound
to H3K9me3 are mostly at the promoter (Fig. 3b–e). MERVL, a TE family
known as a 2C-like marker, bears neither H3K9me3 nor YTHDC1-binding
peaks at its chromatin loci, suggesting that it is regulated indirectly by
YTHDC1 and H3K9me3 (Fig. 3c, Extended Data Fig. 4c). Notably, expres-
sion of YTHDC1–H3K9me3-bound TEs was significantly elevated upon
Ythdc1 KO compared with their YTHDC1–H3K9me3-independent counter-
parts (Extended Data Fig. 4d). Moreover, TE-binding densities of YTHDC1
and H3K9me3 are highly correlated with each other and with the enhance-
ment of TE expression levels upon *Ythdc1* KO (Extended Data Fig. 4e).
These analyses provide evidence supporting the effect of YTHDC1 binding
on H3K9me3-rich chromatin and the subsequent suppression of TEs.

To determine whether YTHDC1 is recruited onto chromatin by
m⁶A-marked TE-derived RNAs, we interrogated the chromatin occu-
pancy of m⁶A-modified TE RNAs using chromatin isolation by RNA puri-
fication (ChIRP-seq). IAP and LINE1 RNAs predominantly target the
chromatin regions of their respective subfamilies (Fig. 3f, Extended Data
Fig. 4f), whose pattern was consistently observed by global RNA inter-
actions with DNA by deep sequencing (GRID-seq) with mouse ES cells[25]

(Extended Data Fig. 4g), fulfilling the prerequisite of YTHDC1-mediated
chromatin targeting on TE regions via their corresponding RNAs.

Since the m⁶A-binding-defective YTHDC1 mutants do not rescue the
Ythdc1 KO-induced 2C-like transition (Fig. 1h, i), we further exam-
ined the m⁶A-dependence of YTHDC1-regulated downstream biologi-
cal consequences by knocking out exons 5 to 7 of *Mettl3* in mouse ES
cells (Extended Data Fig. 5a). Integrating m⁶A RIP-seq data from two
Mettl3-KO mouse ES cell lines (Extended Data Fig. 5b, c) and additional
published datasets[8,14,20], we found that *Mettl3* KO abolishes RNA m⁶A
modification on 25 of the 45 m⁶A-marked TE families in all four datasets
(Extended Data Fig. 5d, e), including IAPA_MM-int, L1Md_A, L1Md_Gf and
L1Md_T (Extended Data Fig. 5e–g). *Mettl3* KO also enhanced the expres-
sion of SETDB1-dependent H3K9me3-modified TEs, largely overlapping
with those that are upregulated upon *Ythdc1* KO (Fig. 3g, Extended
Data Fig. 5h). Consistently, H3K9me3 ChIP-seq further revealed that
global SETDB1-mediated H3K9me3 was impaired upon *Mettl3* KO
(Fig. 3h, Extended Data Fig. 5i), especially on the YTHDC1-dependent
H3K9me3-marked TEs (Fig. 3i, Extended Data Fig. 5j–l). For exam-
ple, both *Mettl3* KO and *Ythdc1* KO dampened H3K9me3 levels of the
RNA-targeting and YTHDC1-targeting TEs IAP and LINE1 (Figs. 2h, 3i,

Article

Extended Data Fig. 5l). ChIRP-seq analysis reveals that the chromatin association of IAP RNA is dependent on YTHDC1 and METTL3, whereas that of LINE1 RNA is independent of YTHDC1 and METTL3 (Extended Data Fig. 5m, n), suggesting that these two TE-derived RNAs bind to chromatin through distinct mechanisms. These results highlight that m6A modification of TE RNAs is required for YTHDC1-mediated chromatin modulation via H3K9me3 and functional regulation through retrotransposon repression.

Consistently, *Mettl3* KO also induces 2C-like genes (Extended Data Fig. 6a), indicating that the 2C-like transition is initiated in an m6A-dependent manner. Of note, LINE1 RNA has been shown to regulate mouse ES cell self-renewal and restrain the 2C-like transition in a *Dux*-dependent manner[23]. *Dux* was upregulated in *Ythdc1*-KO, *Setdb1*-KO and *Mettl3*-KO mouse ES cells, demonstrating that YTHDC1 and m6A are vital for *Dux* regulation (Fig. 3j). ChIRP and GRID-seq further revealed that LINE1 RNA occupied the *Dux* locus (Fig. 3k, Extended Data Fig. 6b), reinforcing the evidence that m6A-modified TE RNA scaffolds YTHDC1 to modulate *Dux* and potentially, 2C-like cell fate. To explore the role of *Dux* in YTHDC1-regulated 2C-like transition, we knocked-out the *Dux* locus in the *Ythdc1*-cKO cells[16,18] (Extended Data Fig. 6c), in which *Dux* was not induced upon *Ythdc1*-KO (Fig. 3l). *Dux* depletion was sufficient to block the induction of the 2C-like program upon *Ythdc1*-KO, as indicated by MERVL reporter and *Zscan4* genes (Fig. 3m, n, Extended Data Fig. 6d), elucidating the predominant role of *Dux* in a 2C-like transition upon loss of *Ythdc1*. Despite the blockage of the 2C-like program (*Zscan4* and MERVL), *Dux*-KO cells retained the ability to reactivate many retrotransposons upon *Ythdc1* depletion (Fig. 3n, Extended Data Fig. 6d) in tandem with dampened H3K9me3, resembling *Dux*-WT cells upon *Ythdc1* KO (Fig. 2h, Extended Data Fig. 6e). In addition, we compared the transcriptomes of *Ythdc1*-KO and *Mettl3*-KO cells; GO terms related to cell cycle were specifically enriched among down-regulated genes in *Ythdc1*-KO cells, but not in *Mettl3*-KO cells (Extended Data Fig. 6f–h). Consistently, early studies have reported difficulties in generating *Ythdc1*-KO mouse ES cells[14,21], even though *Mettl3*-KO mouse ES cell lines have been generated several times independently[8,14,19,20]. These results emphasize that YTHDC1-mediated repression of these retrotransposons is independent of *Dux* and *Dux*-regulated cell-fate decisions.

Discussion

In sum, we report here that YTHDC1 is required to maintain mouse ES cell identity and retrotransposon repression by recognizing a subset of m6A-marked TE-derived transcripts and then recruiting SETDB1. Specifically, SETDB1-mediated H3K9me3 is dependent on YTHDC1 and m6A RNA to repress 2C-like transition and retrotransposons. These results reveal a previously undefined key role of RNA m6A modification and the reader YTHDC1 in chromatin remodelling and early development.

It has been demonstrated that the inhibition of a series of chromatin modifiers, such as the H3K9me3 regulators SETDB1 and KAP1, induces 2C-like transition in mouse ES cells, highlighting their epigenetic restricting roles in 2C-like state and cell-fate decisions[18,26,27]. However, how these regulators are recruited to specific chromatin regions remains unknown. Small RNAs are known to mediate chromatin silencing in fission yeast, plants, *Caenorhabditis elegans*, *Drosophila melanogaster* and ciliates[28], but it remains unclear whether this occurs in mammalian cells. Our study provides evidence that in mouse cells, small RNAs derived from TEs participate in chromatin silencing through m6A RNA modification and the reader YTHDC1. These results provide insights on the role of nuclear small RNAs and RNA modifications in chromatin remodelling. This YTHDC1- and m6A-dependent chromatin remodelling shares a similar mechanism with heterochromatin formation at meiotic chromatin loci mediated by the YTH–containing protein Mmi1 in fission yeast[29]. Nonetheless, Mmi1 targets meiotic mRNAs by recognizing their determinant of selective removal (DSR)

motif through its YTH domain, but in an m6A-independent manner[30], suggesting both evolutionarily conserved and divergent mechanisms of chromatin regulation.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03313-9.

1. Fu, Y., Dominissini, D., Rechavi, G. & He, C. Gene expression regulation mediated through reversible m6A RNA methylation. *Nat. Rev. Genet.* **15**, 293–306 (2014).
2. Zaccara, S., Ries, R. J. & Jaffrey, S. R. Reading, writing and erasing mRNA methylation. *Nat. Rev. Mol. Cell Biol.* **20**, 608–624 (2019).
3. Yang, Y., Hsu, P. J., Chen, Y. S. & Yang, Y. G. Dynamic transcriptomic m6A decoration: writers, erasers, readers and functions in RNA metabolism. *Cell Res.* **28**, 616–624 (2018).
4. Liu, J. et al. YTHDF2/3 are required for somatic reprogramming through different RNA deadenylation pathways. *Cell Rep.* **32**, 108120 (2020).
5. Luo, S. & Tong, L. Molecular basis for the recognition of methylated adenines in RNA by the eukaryotic YTH domain. *Proc. Natl Acad. Sci. USA* **111**, 13834–13839 (2014).
6. Xu, C. et al. Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. *Nat. Chem. Biol.* **10**, 927–929 (2014).
7. Kasowitz, S. D. et al. Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. *PLoS Genet.* **14**, e1007412 (2018).
8. Geula, S. et al. Stem cells. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science* **347**, 1002–1006 (2015).
9. Shi, H. et al. m6A facilitates hippocampus-dependent learning and memory through YTHDF1. *Nature* **563**, 249–253 (2018).
10. Ivanova, I. et al. The RNA m6A reader YTHDF2 is essential for the post-transcriptional regulation of the maternal transcriptome and oocyte competence. *Mol. Cell* **67**, 1059–1067 (2017).
11. Hsu, P. J. et al. Ythdc2 is an *N*6-methyladenosine binding protein that regulates mammalian spermatogenesis. *Cell Res.* **27**, 1115–1127 (2017).
12. Zhang, Y. et al. RNA-binding protein YTHDF3 suppresses interferon-dependent antiviral responses by promoting FOXO3 translation. *Proc. Natl Acad. Sci. USA* **116**, 976–981 (2019).
13. Wu, Y. et al. Nuclear exosome targeting complex core factor Zcchc8 regulates the degradation of LINE1 RNA in early embryos and embryonic stem cells. *Cell Rep.* **29**, 2461–2472 (2019).
14. Liu, J. et al. *N*6-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science* **367**, 580–586 (2020).
15. Macfarlan, T. S. et al. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012).
16. Hendrickson, P. G. et al. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **49**, 925–934 (2017).
17. Genet, M. & Torres-Padilla, M. E. The molecular and cellular features of 2-cell-like cells: a reference guide. *Development* **147**, dev189688 (2020).
18. Wu, K. et al. SETDB1-mediated cell fate transition between 2C-like and pluripotent states. *Cell Rep.* **30**, 25–36 (2020).
19. Batista, P. J. et al. m6A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell* **15**, 707–719 (2014).
20. Chelmicki, T. et al. m6A RNA methylation regulates the fate of endogenous retroviruses. *Nature* https://doi.org/10.1038/s41586-020-03135-1 (2021).
21. Patil, D. P. et al. m6A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* **537**, 369–373 (2016).
22. Jachowicz, J. W. et al. LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat. Genet.* **49**, 1502–1510 (2017).
23. Percharde, M. et al. A LINE1-nucleolin partnership regulates early development and ESC identity. *Cell* **174**, 391–405 (2018).
24. Matsui, T. et al. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* **464**, 927–931 (2010).
25. Li, X. et al. GRID-seq reveals the global RNA-chromatin interactome. *Nat. Biotechnol.* **35**, 940–950 (2017).
26. Iturbide, A. & Torres-Padilla, M. E. A cell in hand is worth two in the embryo: recent advances in 2-cell like cell reprogramming. *Curr. Opin. Genet. Dev.* **64**, 26–30 (2020).
27. Chen, J. et al. H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat. Genet.* **45**, 34–42 (2013).
28. Moazed, D. Small RNAs in transcriptional gene silencing and genome defence. *Nature* **457**, 413–420 (2009).
29. Zofall, M. et al. RNA elimination machinery targeting meiotic mRNAs promotes facultative heterochromatin formation. *Science* **335**, 96–100 (2012).
30. Wang, C. et al. A novel RNA-binding mode of the YTH domain reveals the mechanism for recognition of determinant of selective removal by Mmi1. *Nucleic Acids Res.* **44**, 969–982 (2016).

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Mouse strains

CD-1 mice were used as embryo donors and the pseudopregnant recipients for blastocyst injection. Animal experiments were performed in accordance with the Guide of the Care and Use of Laboratory Animals by the National Research Council and approved by (GIBH) Institutional Animal Care and Use Committee.

Cell lines and cell culture conditions

Male mouse ES cells were derived from 3.5 days post coitum dpc inner cell mass (ICM) from 129 mice (stock no. 217, Beijing Vital River Laboratory Animal Technology). Genetically engineered mouse ES cells, including *Ythdc1*-cKO, Halo-*Ythdc1* knockin (KI), MERVL-tdTomato reporter, *Dux* KO and *Mettl3* KO were constructed and cultured in feeder-free conditions, using serum/LIF medium containing high-glucose Dulbecco's modified Eagle's medium (DMEM, HyClone, SH30022.01), 15% fetal bovine serum (FBS, Front Biomedical, OPT500), 1× GlutaMAX, 1× nonessential amino acids (NEAA, Gibco, 11140076), 1× penicillin-streptomycin (Hyclone, SV30010), 1× sodium pyruvate (Gibco, 11360070), 1 mM 2-mercaptoethanol (Gibco, 21985023) and 1,000 U ml^{-1} LIF. HEK 293T (ATCC, CRL-1126) cells were maintained in high-glucose DMEM supplemented with 10% FBS (Natocor, SFBE), 1× GlutaMAX (Gibco, 35050079) and 1× NEAA. All cell lines were cultured at 37 °C in a humidified atmosphere containing 5% CO_2. Mycoplasma detection tests were conducted routinely to ensure mycoplasma-free conditions throughout the study.

Generation of gene-editing cell lines

We applied CRISPR–Cas9 system to construct the gene-editing cell lines in male 129 mouse ES cells. The specific guide RNA (gRNA) sequences targeting genome were designed and inserted into the pX330 plasmid (Addgene, 42230). For generation of *Ythdc1*-cKO mouse ES cells, a *Ythdc1* targeting vector was assembled by inserting two loxP sites at introns 6 and 9, as well as a FRT-flanked PGK-Puro cassette after intron 9 and before the second loxP site, and the PGK-CreERT-Puro cassettes were introduced into the other allele, replacing exons 7–9 of *Ythdc1* by homologous recombination. For construction of Halo-*Ythdc1*-KI mouse ES cells, The targeting vector contains the Halo tag (a gift from Y. Yu, Institute of Biophysics, Chinese Academy of Sciences), following three tandem copies of TEV cleavage sites, and two copies of StrepII sequences tightly after the translation initiation codon, ATG. In addition, the FRT-flanked PGK-Puro cassette was inserted into the first intron. For generation of Rosa26 locus knock-in 2C-tdTomato reporter cell line on *Ythdc1*-cKO mouse ES cells, the Rosa26-targeting MERVL-LTR-tdTomato reporter (a gift from Y. Huang, Chinese Academy of Medical Sciences) was used. For generation of *Dux* KO in *Ythdc1*-cKO mouse ES cells, the *Dux* KO-targeting construct was constructed as previously described[31], and the *Dux* cluster was deleted by replacing resistance genes in *Ythdc1*-cKO mouse ES cell line with MERVL-tdTomato reporter. For generation of *Mettl3*-KO mouse ES cells, 129 mouse ES cells were targeted for disruption of the endogenous *Mettl3* locus using the CRISPR–Cas9 system and homologous recombination. Exons 5, 6 and 7 of mouse *Mettl3* were deleted on the basis that they contain the DPPW active motif. Next, the gRNAs and the linearized targeting vector were electroporated into the mouse ES cells using the Mouse Embryonic Stem Cell Nucleofector Kit (Lonza, VPH-1001). After screening (1 µg ml^{-1} puromycin, 400 µg ml^{-1} G418 or 200 µg ml^{-1} hygromycin) for 3–5 days, colonies were picked and validated by genotyping PCR or western blot. Furthermore, the PGK-Puro in *Ythdc1*-cKO mouse ES cells and Halo-*Ythdc1*-KI mouse ES cells was excised by transient transfection with pCAG-Flp plasmid using Lipofectamine 3000 Transfection Reagent (Invitrogen, L3000015), and *Ythdc1*-KO was initiated with 1 µM 4OHT (Sigma, H7904) treatment for 4 days.

Plasmid construction, cell transfection and lentivirus production

Mouse *Ythdc1*, *Ythdc1*-mut, *Ythdc1*-W378A, *Ythdc1*-W429A and *Setdb1* were cloned into the pRlenti-3× Flag vector. *Ythdc1* and *Setdb1* were also cloned into the pRlenti-HA vector. For transient expression, cells were transfected by Lipofectamine 3000 Transfection Reagent (Invitrogen, L3000015) according to the manufacturer's instructions. For lentivirus generation, HEK 293T cells were plated at density of 1.5×10^6 cells per well in a 6-well plate and cultured overnight. When the confluency was around 80%, pRlenti vector containing target genes was co-transfected with the packaging plasmids psPAX2 and pMD.2G using the standard calcium phosphate method. Cells were then incubated overnight and the medium was changed after 12 h. The viral supernatants were collected 48 h and 72 h post-transfection. After filtration through a 0.45-mm filter (Millipore), the supernatants were stored at 4 °C and used within 24 h.

Immunoblotting

Cells were digested with 0.25% trypsin (Gibco, 25200114), lysed in the RIPA buffer supplemented with EDTA-free protease inhibitor cocktail (Roche, 4693132001) on ice for 15 min, and then boiled at 100 °C for 10 min. After centrifugation (14,000*g* for 10 min at 4 °C), the supernatants were fractionated by SDS–polyacrylamide gel electrophoresis and blotted onto PVDF membranes (Millipore, IPVH00010). The membranes were incubated with primary antibodies at room temperature for 2 h or 4 °C overnight after first blocking with skimmed milk. After washing with PBST, the membranes were incubated with secondary antibodies conjugated to HRP at room temperature for 1 h. ECL solution (Millipore) was added and signals were detected using a MiniChemi 610 Chemiluminescence Imaging System (Beijing Sage Creation). Protein quantification was performed using the Enhanced BCA Protein Assay Kit (Beyotime, P0010) according to the manufacturer's instructions. The raw uncropped images are available in Supplementary Fig. 1.

Cell viability assay

WT and *Ythdc1*-cKO ES cells were pre-treated with DMSO or 1 µM 4OHT for 4 days, digested with 0.25% trypsin, and plated onto 96-well plates (5,000 cells per well) with 100 µl serum/LIF medium containing DMSO or 1 µM 4OHT . After 24 h, 48 h, 72 h or 96 h, cell viability was detected using the Cell Counting Kit-8 (CCK8, Beyotime, C0039) according to the manufacturer's instructions. In brief, 10 µl CCK8 solution was added to each well and incubated for 1 h, and the absorbance at 450 nm was determined using a spectrophotometer. Wells with no cells were used as control.

Apoptosis detection

Apoptosis detection was performed using the PE-Annexin V Apoptosis Detection Kit I (BD Biosciences, 559763) according to the manufacturer's instructions. In brief, *Ythdc1*-cKO mouse ES cells were plated onto 6-well dishes and treated with 1 µM 4OHT or DMSO for 4 days before collection before passage or 24 h after passage. The cells were washed twice with cold PBS and resuspended in 1× annexin V binding buffer at a concentration of 1×10^6 cells per ml. Five microlitres of PE-Annexin V was added into 100 µl cell suspension (1×10^5 cells) and incubated for 15 min at room temperature in the dark. Subsequently, 400 µl 1× annexin V binding buffer to each sample and analysed by flow cytometry within 1 h.

EdU assay

Cell proliferation assay was performed using the BeyoClickTM EdU Cell Proliferation Kit with Alexa Fluor 647 (Beyotime, C0081). The cells were

Article

treated with 1 μM 4OHT or DMSO for 3.5 days and seeded in 96-well plates at a density of 3×10^3 cells per well for 24 h. Then, the cells were incubated with 10 μM EdU for 10 min at 37 °C. After being fixed with 4% paraformaldehyde for 15 min, the cells were treated with 0.3% Triton X-100 for 15 min and rinsed three times with 3% BSA. Thereafter, the cells were exposed to 50 μl of click reaction cocktail for 30 min at room temperature, rinsed three times with 3% BSA and then incubated with DAPI to stain nuclei. Images were captured using Olympus IXplore SpinSR microscope. All the experiments were performed in duplicate.

Ythdc1-cKO rescue experiment
WT and *Ythdc1* cKO ES cells were plated onto 6-well dishes at a density of 2×10^5 cells per well. Serum/LIF medium supplemented with 1 μM 4OHT, 4 μg ml^{-1} polybrene, and WT or mutant *Ythdc1* viral supernatants were simultaneously added into the corresponding wells. After 24 h of infection, fresh serum/LIF medium supplemented with 1 μM 4OHT was replaced. Four days later, cells were digested with 0.25% trypsin and replated into 12-well plates with 1×10^5 cells per well. The remaining cells were collected for western blot analysis and RT–qPCR analysis. After treatment with 1 μM 4OHT for another 4 days, images were obtained by phase-contrast microscopy and cell numbers were counted after trypsin digestion.

RNA extraction and RT–qPCR analysis
To extract the total RNAs, cells were collected and lysed with 1 ml TRIzol reagent (MRC, TR118-200), followed by addition of 200 μl chloroform. After centrifugation (14,000*g* for 10 min at 4 °C), the total RNAs in the supernatant were precipitated with isopropanol. cDNAs were reverse transcribed with oligo-dT primers (TaKaRa, 3806) and ReverTra Ace (Toyobo, TRT-101). RT–qPCR was performed on the CFX96 system (Bio-Rad) using ChamQ Universal SYBR qPCR master mix (Vazyme, Q711-02). Gene expression levels were normalized to those of *Gapdh*. Primers used in this study are shown in Supplementary Table.

Fluorescence-activated cell sorting
Ythdc1-cKO MERVL::tdTomato reporter mouse ES cells were seeded onto 6-well plates and cultivated in serum/LIF medium supplemented with DMSO or 1 μM 4OHT. After 4 days of treatment, cells were digested with 0.25% trypsin and resuspended in PBS with 1% FBS. The cell suspension was filtered with a strainer and analysed using a LSRFortessa flow cytometer (BD Biosciences). The tdTomato intensity was detected in the PE channel. Data were analysed using FlowJo v.10 software. For sorting, the cells were resuspension with flow cytometry buffer (PBS with 2% FBS) and sorted with Arial (BD Biosciences) to separate tdTomato-positive and -negative mouse ES cells.

Chimeric blastocyst assay
Chimeric embryo assay was performed by the Turbo Mice Core of Cell Lineage and Atlas Technology Facilities. The zona pellucida of the 8-cell stage embryos were removed by a short exposure to acidic Tyrode's solution (Sigma, T1788). Two denuded embryos were placed into each of the concaved microwells created by an aggregation needle. Ten to fifteen ES cells with a CAG-mCherry transgene were transferred into each concaved microwell and cocultured with the denuded embryos. Before and after the aggregation procedure, all embryos were cultured in potassium simplex optimized medium (KSOM) medium[32]. To obtain E4.5 chimeric embryos for observing contribution of ES cells and immunostaining, the aggregated embryos were cultured at 37 °C in 5% CO_2 for 48 h. To obtain E6.5 embryos, embryos were transferred into the uterine horns of 2.5 dpc pseudopregnant CD1 mice and collected 4 days later.

Immunofluorescence staining
Cells growing on a glass-bottom dish (801002, NEST) were fixed in 4% paraformaldehyde for 30 min, washed with PBS three times and permeabilized with 0.2% TritonX-100 and 1% BSA for 30 min at room temperature. The cells were then incubated with anti-NANOG antibody (A300-397A, Bethyl) for 2 h. Following PBS washes, secondary antibody was applied for 1 h and subsequently, DAPI was applied for 2 min. Finally, coverslips were mounted on slides for observation.

Whole-mount immunofluorescence staining of E4.5 or E6.5 embryos was performed as previously described[33]. In brief, embryos were fixed in 4% paraformaldehyde at room temperature for 30 min, washed 3 times in PBS-T (PBS +0.1% Triton X-100), and permeabilized for 1 h in PBS 0.5% Triton X-100. E4.5 embryos were blocked in PBS-T with 3% BSA at room temperature for 1 h and E6.5 embryos were blocked in PBS-T with 1% BSA (Sigma) and 3% normal donkey serum (Jackson ImmunoResearch Laboratories) overnight at 4 °C. Primary and secondary antibodies were diluted in the blocking solution. E4.5 and E6.5 embryos were incubated with primary antibodies at 4 °C overnight or 48 h, respectively. The following primary antibodies were used: mouse anti-CDX2 (1:100; BioGenex, MU392A-UC), anti-mCherry (1:400; LSBio, LS-C204207), anti-ELF5 (1:200; Santa-Cruz, sc-166653). After three washes in PBS-T, the embryos were incubated with secondary antibodies. Finally, the embryos were washed 4 times in PBS-T and stained in 10 μg ml^{-1} DAPI for 1 h and mounted in PBS in a glass-bottom dish for observation. The Dragonfly 200 spinning disk confocal microscope (Andor Technology) was operated with a 60× oil lens with 405, 488 and 568 nm lasers and with the Turbo Mice Core of Cell Lineage and Atlas Facility. The raw high-resolution images are available in Supplementary Fig. 2.

Co-immunoprecipitation
HEK 293T cells were plated onto a 6-cm dish and co-transfected with Flag- and HA-tag plasmids encoding YTHDC1 and SETDB1. Transfected cells were collected after 48 h and homogenized in 1 ml lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 2 mM EDTA, 1% NP-40 and 5 μg ml^{-1} of RNase A (Thermo Scientific, EN0531)) supplemented with EDTA-free protease inhibitor cocktail (Roche, 4693132001) for 10 min on ice. Cells were further lysed by sonication with low energy and the supernatant was collected after 14,000*g* centrifugation for 10 min at 4 °C. After taking 80 μl as the input, 400 μl was used for immuno-precipitation. The immunoprecipitation sample was incubated with 2 μg anti-Flag antibody (Sigma, F1804). Two micrograms of mouse IgG (Abcam, ab18413) were applied as a negative control. Next, 50 μl of protein A and protein G Dynabeads mixture (GE Healthcare) was added to the samples and incubated overnight at 4 °C with gentle rotation. Beads were repeatedly washed 5 times using 1 ml lysis buffer for 5 min at 4 °C. The protein complex was eluted by boiling at 100 °C for 10 min in 100 μl of 1× loading buffer and analysed by immunoblotting. Antibodies used in this study are listed in Supplementary Table.

RNA-seq and data analysis
Total RNAs were extracted as described above. The Total RNA-seq (H/M/R) Library Prep Kit for Illumina (VAHTS, NR603) was used for RNA library preparation. In brief, 1 μg total RNAs were hybridized with the rRNA probe (H/M/R) and digested by RNase H to remove ribosomal RNAs. After DNase I digestion, the ribosomal-depleted RNAs were fragmented at 94 °C for 8 min. Then, the first-strand and second-strand cDNAs were synthesized successively using the provided reagents. The cDNA was purified by VAHTS DNA Clean Beads (VAHTS, NR411), followed by end repair, dA-tailing, adaptor ligation and second-strand cDNA digestion. After two rounds of purification, the cDNAs were PCR-amplified and purified using VAHTS DNA Clean Beads.

The RNA-seq data processing was performed as described[34]. To analyse the gene expression, reads were aligned to the reference transcriptome using RSEM[35] and bowtie2[36] (–bowtie2 –bowtie2-sensitivity-level very_sensitive –no-bam-output –estimate -rspd) and the index built by RSEM with the mouse genome, mm10, and Enemble gene annotation track v.74. For TE-expression analysis, the reads were mapped to the mouse genome mm10 using the STAR aligner[37] and the counts for each

gene or TE family were calculated using the TEtranscripts[38]. DESeq2[39] was used for data normalization and differential expression analysis. Differentially expressed genes and TEs were defined by a Benjamini–Hochberg-corrected *P* value <0.05 and an absolute fold change >2 (for genes) or an absolute fold-change >1.5 (for TEs).

ChIP-seq

SETDB1 and YTHDC1 ChIP was performed as described previously with minimal change[27]. In brief, chromatin samples were incubated with specific antibodies in ChIP lysis buffer (20 mM Tris-HCl pH8.1, 150 mM NaCl, 2 mM EDTA, 1% TritonX-100 and 0.05% SDS) overnight at 4 °C. The elution and reverse cross-linking processes were carried out in elution buffer (50 mM Tris-HCl pH8.0 and 1% SDS) at 65 °C for 5 h. After 1 h of RNase A (1 U μl^{-1}) treatment at 37 °C and Proteinase K (1 U μl^{-1}) digestion at 55 °C, DNA samples were then purified using a PCR extraction kit (QIAGEN 28006). H3K9me3 ChIP-seq was performed as described[40]. In brief, 1×10^7 cells were collected and resuspended in 500 μl buffer 1 (0.32 M sucrose, 15 mM Tris-HCl pH 7.5, 60 mM KCl, 15 mM NaCl, 5 mM MgCl$_2$, 0.5 mM DTT, 0.1 mM EGTA, 0.1 mM PMSF and 1:1,000 EDTA-free protease inhibitor cocktail (Roche, 4693132001)) supplemented with 0.1% IGEPAL CA-630 (Sigma, I8896) on ice for 10 min. The lysate was carefully layered on top of 1 ml buffer 3 (1.2 M sucrose, 15 mM Tris-HCl pH7.5, 60 mM KCl, 15 mM NaCl, 5 mM MgCl$_2$, 0.5 mM DTT, 0.1 mM EGTA, 0.1 mM PMSF, 1:1,000 protease inhibitor cocktail). After centrifugation (10,000*g* for 20 min at 4 °C), nuclei were resuspended in 500 μl buffer A (0.34 M sucrose, 60 mM KCl, 15 mM HEPES pH7.4, 15 mM NaCl, 4 mM MgCl$_2$, 1 mM DTT, 0.1 mM PMSF, 1:1,000 protease inhibitor cocktail) supplemented with 3 mM CaCl$_2$ and 2 U μl^{-1} MNase (NEB, M0247S) and digested for 15 min at 37 °C. Then, 5 mM EGTA was added to stop the reaction. After centrifugation at 13,500*g* for 10 min at 4 °C, chromatin was resuspended in 10 mM EDTA (pH 8.0) supplemented with 1 mM PMSF and 1:1,000 protease inhibitor cocktail, followed by rotation at 4 °C for 2-4 h. The mixture was adjusted to 500 mM NaCl and rotated for additional 45 min. After centrifugation, the supernatant containing chromatin was diluted to 100 ng/μl with buffer B (20 mM Tris-HCl pH8.0, 5 mM EDTA, 500 mM NaCl, 0.2% Tween 20) and the fragmentation of chromatin was detected by agarose electrophoresis. A total of 2.5 μg fragmented chromatin was kept as the input. For immunoprecipitation, the reaction mixture, including 20 μg fragmented chromatin, 2 μg anti-H3K9me3 antibody (Abcam, ab8898), and 50 μl protein A/G dynabeads (GE Healthcare) was incubated overnight with gentle rotation at 4 °C. The beads were washed three times with buffer B, and once with buffer B without Tween 20 at 4 °C for 5 min. The DNA was eluted with 300 μl elution buffer (20 mM Tris-HCl pH8.0, 20 mM EDTA, 0.5% SDS), digested with proteinase K enzyme (NEB, P8107S), and purified using a MinElute kit (Qiagen, 28006). The ChIP DNA library was constructed with a KAPA HTP/LTP Library Preparation Kit according to the manufacturer's instructions.

RNA immunoprecipitation

The YTHDC1 RIP assay was performed based on GoldCLIP[41] with modifications. Cells expressing Halo–GFP or Halo–YTHDC1 were collected, lysed, and pulled-down using Halo-tagged beads (Promega, G7282) as previously described[41], except that no nuclease digestion or RNA fragmentation conducted throughout the experiment. The binding efficiency of Halo-tagged protein to the beads was evaluated by western blot. After the denaturing wash, the RNAs were released from the proteins by proteinase K enzyme (NEB, P8107S) digestion. For RIP-RT–qPCR, the equivalent input and immunoprecipitated RNAs were reverse-transcribed by SuperScript III Reverse Transcriptase (Invitrogen, 18080044) using random hexamers, followed by qPCR to quantify the transcripts. The enrichment score of a transcript was the ratio of the quantified immunoprecipitated fraction to its total input, using GFP as a standard for normalization. For RIP-seq, RNAs from the input and immunoprecipitated samples were used to generate libraries using QIAseq miRNA Library Kit (Qiagen, 331502), according to the manufacturer's instructions.

For m^{6}A RIP-seq, PolyA-tailed mRNA was purified using the Dynabeads mRNA purification kit (Thermo Fisher). To perform methylated RNA immunoprecipitation (MeRIP), purified mRNA was first fragmented to ~200 nucleotides (nt) in length using RNA Fragmentation Reagent (Thermo Fisher). External spike-in oligonucleotides were also added to the RNA pool for quality control. Then, immunoprecipitation was conducted following the protocol from EpiMark N^6-Methyladenosine Enrichment kit (NEB, E1610S). Pull-down RNA fragments were eluted with 130 μl elution buffer (150 mM NaCl, 10 mM pH 7.4 Tris-HCl, 0.1% Igepal CA-630, 6.7 mM m^{6}A in nuclease-free H$_2$O) at 37 °C and collected by ethanol precipitation. The immunoprecipitation products and input sample from before immunoprecipitation were used to construct NGS libraries using NEBNext Ultra II Directional RNA Library Prep (New England Biolabs, 625 E7760S).

Chromatin isolation by RNA purification

ChIRP was performed as previously described with some modifications[23,42]. In brief, a set of biotinylated antisense oligonucleotides with 59-nt-long probes tiled along the L1 consensus sequence or *Malat1* RNA, and 20-nt probes tiled along the IAP consensus sequence or LacZ. Probes were labelled on the basis of their positions along the RNA, and separated them into 'even' and 'odd' pools, to create two independent probe sets for ChIRP experiments. *Ythdc1*-cKO ES cells were treated with DMSO or 1 μM 4OHT for 4 days; cells were then collected by 0.25% trypsin digestion and cross-linked with 2 mM DSP (Thermo Scientific, 22586) in PBS at room temperature for 30 min with gentle end-to-end rotation. The cells were further fixed in suspension for 10 min in 3.7% formaldehyde, and quenched with 250 mM glycine for 5 min at room temperature. Cells were then washed with ice-cold PBS for three times and snap frozen in liquid nitrogen. 1×10^7 cells were resuspended with 500 μl DNase I digestion solution (20 mM Tris-HCl, pH7.5, 5 mM MgCl$_2$, 0.5 mM CaCl$_2$, 0.5% Triton X-100) with 1/20 volume of ribonucleoside vanadyl complex (RVC, NEB, S1402S), 5 μl protease inhibitor cocktail (Roche, 4693132001), 5 μl 100 mM PMSF and 12 U ml^{-1} Dnase I (Invitrogen, AM2222); the reaction was rotated for 10 min at 37 °C and stopped with 20 mM EDTA. After centrifugation, chromatin was resuspended in nuclear lysis buffer (50 mM Tris-HCl, pH7.5, 10 mM EDTA, 1% SDS, protease and RNase inhibitors) and sonicated with Bioruptor Pico (5 cycles, 30 s on, 30 s off, Diagenode). The supernatant was collected for ChIRP experiments after centrifuged at 14,000*g* for 10 min at 4 °C and a sample was saved for use as the 'input' sample. For the Rnase treatment control, samples were treated with 10 μg ml^{-1} Rnase A/T1 Mix (Thermo Scientific, EN0551) at 37 °C for 20 min. For hybridization, samples were incubated with ChIRP probes (1 μl of 100 μM probes per 1 ml of lysate), supplemented with one-fourth volume of 5× hybridization buffer (50 mM Tris-HCl, pH7.5, 10 mM EDTA, 1.5 M NaCl, 50% formamide). After rotation at 39 °C for 3 h, 50 μl prebalanced streptavidin M280 beads were then added into the hybridization and incubated for an additional 3 h. The beads were washed 5 times with 0.2× SSC wash buffer with 1% SDS at 42 °C. For elution, the beads were washed once with SDS elution buffer (50 mM Tris-HCl, 5 mM MgCl$_2$, 75 mM NaCl, 1% SDS) at 39 °C for 20 min, and once with elution buffer (50 mM Tris-HCl, 5 mM MgCl$_2$, 75 mM NaCl, 0.1% Triton X-100) at 39 °C for 5 min. DNA was eluted from the beads by RNase H treatment in two sequential incubations with Rnase H (NEB, M0297L) at 37 °C for 20 min, and with SDS elution buffer at room temperature for 2 min, combining all eluents. Cross-linking was reversed by treatment with 0.1 μg μl^{-1} protease K (NEB, P8107S), 150 mM NaCl and 10 mM EDTA at 65 °C overnight and the DNA was purified using the MinElute PCR Purification Kit (Qiagen, 28006). ChIRP enrichments were analysed by qPCR of the purified DNA. The libraries were constructed by a KAPA HTP/LTP Library Preparation Kit (KAPA, KK8234) according to the manufacturer's instructions. Probes used in this study are available in the Supplementary Table.

Article

4sUDRB-seq

Metabolic labelling nascent RNA with 4sU and nascent RNA purification was performed as described[43]. In brief, *Ythdc1*-cKO ES cells were pre-treated with DMSO or 1 μM 4OHT for 4 days, then incubated with 100 μM DRB (Sigma, D1916) for 3 h to block RNA Pol II transcription. Cells were washed 3 times with PBS to remove DRB and recover transcription, and newly transcribed RNA was labelled with 1 mM 4sU (Sigma, T4509) for 0, 4 and 8 min. Transcription was stopped by adding TRizol (Invitrogen) into the cells and total RNA was extracted using the miRNeasy mini kit (Qiagen, 217004) with the RNase-free DNase set (Qiagen, 79254). One-hundred micrograms total RNA was used for biotinylation and purification of 4sU-labelled nascent RNA. The 4sU-labelled RNA was incubated with biotin HPDP–DMF stock (50 mg of EZ-Link HPDP-biotin (Thermo Scientific, 21341) dissolved in 50 ml DMF (Sigma, D4551)) in biotinylation buffer (10 mM Tris pH 7.4, 1 mM EDTA) for 2 h on a tube rotator at room temperature. Unbound HPDP-biotin was removed by equal volume chloroform extraction twice, and RNA was precipitated at 14,000 rpm for 30 min at 4 °C with 1:10 volume of 5 M NaCl and an equal volume of isopropanol. The RNA pellet was washed twice with 75% ethanol and resuspended in 100 μl nuclease-free water. Biotinylated RNA was enriched by 100 μl streptavidin T1 (Invitrogen, 65601) at room temperature for 20 min. Beads were washed 5 times with 1 ml washing buffer (100 mM Tris pH 7.4, 10 mM EDTA, 1 M NaCl, 0.1% Tween 20). Nascent RNA was eluted with 100 μl of 100 mM dithiotheitol (DTT, Thermo Scientific, A39255) twice and was purified with the RNeasy MinElute clean-up kit (Qiagen, 74204) according to the manufacturer's instructions.

ChIP-seq data analysis

ChIP-seq reads were aligned to the mouse mm10 genome using bowtie2[36] with the options '-p 20 –very-sensitive –end-to-end –no-unal –no-mixed -X 2000'. The multimapped reads were kept for TE analysis, but only the best alignment was reported for those reads; if more than one equivalent best alignment was found, then one random alignment was reported. Reads mapping to mitochondrial DNA or unassigned sequences were discarded. For unique alignments, duplicate reads and low-mapping-quality reads were filtered using SAMtools (v.1.9). ChIP-seq enriched peaks were called by MACS2 (v.2.1.2)[44], and the common peaks were merged using bedtools and used for subsequent analysis. The enriched peaks were annotated using the annotatePeaks. pl script from the Homer package[45]. For visualization, BAM files of biological replicates were merged using SAMtools (v.1.9)[46]. BigWig files were generated using deeptools[47] with the RPKM normalization method. Figures illustrating these continuous tag counts over selected genomic intervals were created in the IGV browser[48]. The H3K9me3 ChIP-seq data of Setdb1-KO mouse ES cells was obtained from the BioProject accession PRJNA544540.

The ChIP-seq signal for TEs was processed as previously described[34]. The coordinates and annotations of TEs were downloaded from the RepeatMasker (http://www.repeatmasker.org). As many ChIP-seq signals were not evenly distributed across TEs—for example, YTDHC1–H3K9me3 binding prefers 5′ untranslated region (UTR) of intact L1 s—we divided the TEs into evenly spaced 500-bp bins, the coverage signal on each bin was extracted using deeptools[47], and the bin with the maximum signal was used as the observed immunoprecipitate value, and then divided by normalized read counts of a matched input sample and expressed as $\log_2$ fold enrichment (Fig. 3c). In addition, the following quality controls were performed: read distribution across the repetitive sequence was inspected using IGV genome browser for each repeat family to confirm coverage of the whole repetitive sequence. To avoid over- or under-estimating fold enrichments due to low coverage of sequencing reads, random genomic regions of the same number as the TE type and the same size as the TE were taken, the coverage signal was measured, and the maximal bin was taken as the randomly expected background value. TEs with observed IP values smaller than the randomly expected background value were excluded from analysis.

RIP-seq data analysis

Raw reads were trimmed with TrimGalore (v.0.6.4) (https://github.com/FelixKrueger/TrimGalore), then aligned to the mouse genome (mm10) and transcriptome (GENCODE version M21) using HISAT2 (v.2.2.0)[49]. The multimapped reads were kept for TE analysis, but only the best alignment was reported for those reads, if more than one equivalent best alignment was found, then one random alignment was reported. Reads mapping to mitochondrial DNA or unassigned sequences were discarded. For unique alignments, duplicate reads low mapping quality reads were filtered by SAMtools[46]. The BAM file were converted to Bigwig format, and the read coverage was normalized to the library size for input and eluate (m6A–YTDHC1 RIP) respectively by RPKM normalization using deeptools[47]. The average normalized read densities of merged replicates are shown in the figures.

To search for enriched peaks in the RIP samples compared to the input controls, we adopted the strategy described[50]. We scanned the genome using a 100-nt sliding window with a 50-nt overlap. The mean coverage for each window was calculated for the RIP and input control samples respectively. For each window as assigned, the enrichment fold change of RIP sample over input control sample for the enriched window was defined as at least fourfold over the input control, and the overlapping enriched windows were merged as peaks. To determine high-confidence peaks, common peaks in all the biological replicates were merged and used for downstream analysis.

Quality control was performed by measuring the RPKM at all high-confidence peaks, and then combining biological replicates where the Pearson correlation was >0.8. Outlier samples that did not correlate closely (Pearson correlation was <0.6) with other RIP-seq data samples of the same type from other independent studies were deleted in the downstream analysis. To study the peak distribution across transcripts, we assigned each high-confidence peak to the 'metagene' transcripts, the 5′ UTR, conding sequence and 3′UTR region from Ensemble (v.74) for each gene, and the peak density across all mRNA genes was calculated (Extended Data Fig. 2b). The RIP peak annotation and motif finding were performed by annotatePeaks.pl and findMotifsGenome.pl scripts from Homer. For *Mettl3* knockout m6A RIP-seq analysis, the reads were normalized by m6A spike-in, with the number of reads mapped to the genome divided by number of reads mapped to m6A modified spike-in RNA (Gaussia Luciferase) represented whole m6A level. For RIP-seq TE analysis basically similar to the ChIP-seq data, the enrichment score was calculate using RIP/INPUT, and only the signal consistent in all biological replicates were kept.

ChIRP-seq data analysis

Raw reads were aligned to mouse genome (mm10) using bowtie2 with the parameters "–very-sensitive –end-to-end –no-unal –phred33 – no-mixed -X 2000 –gbar 200". The multimapped reads were kept for TE analysis, but only the best alignment is reported for those reads, if more than one equivalent best alignment was found, then one random alignment was reported. BigWig files were generated using deeptools with the RPKM normalization method. Positive peaks were identified with DFilter[51] program with parameters "-ks=100 -lpval=5". In order to eliminate false-positive hits, only the overlapped signal from both "even" and "odd" probes were considered. And then the Bam files of "even" and "odd" probes were merged using SAMtools (v.1.9) for visualization. Other TE analysis is the same to ChIP-seq data analysis.

4sUDRB-seq data analysis

Raw reads were aligned to mouse genome (mm10) using bowtie2 with the default parameters (only set -N 1)[43]. For each experiment we generated a genome-wide profile of the counts per disjointed 100-bp adjacent bins. For gene or TE quantification, the sum of all bins within genes (exon and intron) or TEs was used. The counts of each experiment was

normalized to rRNA counts. All gene annotations were taken from Ensemble and TE annotations were from Repeatmarsker.

GRID-seq data analysis

The raw reads were downloaded from accession GSE82312[25], and the RNA and DNA reads were separately aligned to mouse genome (mm10) using bowtie2[36] with default parameters. The multimapped reads were kept for TE analysis, but only the best alignment is reported for those reads, if more than one equivalent best alignment was found, then one random alignment was reported. Then the DNA reads and their RNA read mates were mated by custom script, the unmated reads were filtered out for the downstream analysis. To identify IAP/LINE1 RNA-chromatin interaction from GRID-seq, we first summarized the coverage of DNA reads mated with each RNA in 1-Kb binned genome, and then normalized by the total number of mapped reads and the length of chromosome where the bin was located. Only the RNA–DNA interactions detected in both two biological replicates were kept, and the biological replicates were merged for visualization in the figures.

Statistical analysis

The experiments in this study were performed with at least three biological replicates unless specified. Data are presented as the mean ± s.d. or mean ± s.e.m. calculated by GraphPad Prism 6.0 software. Two-tailed unpaired Student's t-test, Mann–Whitney U test and Fisher's exact test were performed in the Graphpad Prism 6.0 and R Bioconductor, respectively. Statistical analysis was performed as described in the figure legends and P-values <0.05 were considered statistically significant. All statistical tests, resulting P values and numbers of observations are indicated in the figure panels or legends.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The data supporting the conclusions of this Article, including H3K9me3–YTHDC1 ChIP–seq, m6A–YTHDC1 RIP-seq, ChIRP-seq, 4sUDRB-seq and RNA-seq data are available at GEO under accession GSE146467. The m6A RIP-seq data were from GSE52662[19], GSE61998[8], GSE145315[20] and GSE133599[14]. The H3K9me3 ChIP-seq data of *Setdb1*-KO mouse ES cells was obtained from the BioProject accession PRJNA544540[18]. The raw uncropped data for gels are appended in Supplementary Fig. 1, and high-resolution images for whole-mount fluorescence imaging are available at Supplementary Fig. 2. qPCR primers, ChIRP probes and antibodies used in this study are listed in the supplementary tables.

31. Chen, Z. & Zhang, Y. Loss of DUX causes minor defects in zygotic genome activation and is compatible with mouse development. *Nat. Genet.* **51**, 947–951 (2019).
32. Summers, M. C., McGinnis, L. K., Lawitts, J. A., Raffin, M. & Biggers, J. D. IVF of mouse ova in a simplex optimized medium supplemented with amino acids. *Hum. Reprod.* **15**, 1791–1801 (2000).
33. Wu, G. et al. Establishment of totipotency does not depend on Oct4A. *Nat. Cell Biol.* **15**, 1089–1097 (2013).
34. He, J. et al. Transposable elements are regulated by context-specific patterns of chromatin marks in mouse embryonic stem cells. *Nat. Commun.* **10**, 34 (2019).
35. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
36. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
37. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
38. Jin, Y., Tam, O. H., Paniagua, E. & Hammell, M. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* **31**, 3593–3599 (2015).
39. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
40. Hasson, D. et al. The octamer is the major form of CENP-A nucleosomes at human centromeres. *Nat. Struct. Mol. Biol.* **20**, 687–695 (2013).
41. Gu, J. et al. GoldCLIP: Gel-omitted Ligation-dependent CLIP. *Genomics Proteomics Bioinformatics* **16**, 136–143 (2018).
42. Chu, C., Qu, K., Zhong, F. L., Artandi, S. E. & Chang, H. Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA–chromatin interactions. *Mol. Cell* **44**, 667–678 (2011).
43. Fuchs, G. et al. 4sUDRB-seq: measuring genomewide transcriptional elongation rates and initiation frequencies within cells. *Genome Biol.* **15**, R69 (2014).
44. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. *Nat. Protocols* **7**, 1728–1740 (2012).
45. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
46. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
47. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44** (W1), W160–W165 (2016).
48. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
49. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
50. Dominissini, D. et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201–206 (2012).
51. Kumar, V. et al. Uniform, optimal signal processing of mapped deep-sequencing data. *Nat. Biotechnol.* **31**, 615–622 (2013).
52. Ng, R. K. et al. Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. *Nat. Cell Biol.* **10**, 1280–1290 (2008).
53. Latos, P. A. & Hemberger, M. From the stem of the placental tree: trophoblast stem cells and their progeny. *Development* **143**, 3650–3660 (2016).

Author contributions J.L. and M.G. performed the main experiments; J.H. performed the bioinformatics analysis. J.L. and J.C. initiated the study. K.W. performed the chimeric embryo assay, 2C reporter cell lines construction and *Dux* knockout. S.L., L.J. and Y.C. assisted in experiments including with cell culture, plasmid construction and imaging. G.W., M.Z. and K.W. performed chimeric embryo immunofluorescence staining assays and L.C. assisted embryo imaging. H.L. performed high-throughput sequencing. J.S. assisted with bioinformatics analysis. J.L. performed YTHDC1 RIP-seq, ChIRP-seq and 4sUDRB-seq with help from X.W., Y.L. and X.B. M.G. performed *Mettl3* KO and Y.-L.Z. and G.-Z.L. performed the m6A RIP-seq. G.W. and X.Z. contributed to the work. J.C. wrote the manuscript and D.P., J.W., J.L. and J.H. helped to improve it. J.C. conceived and supervised the study.
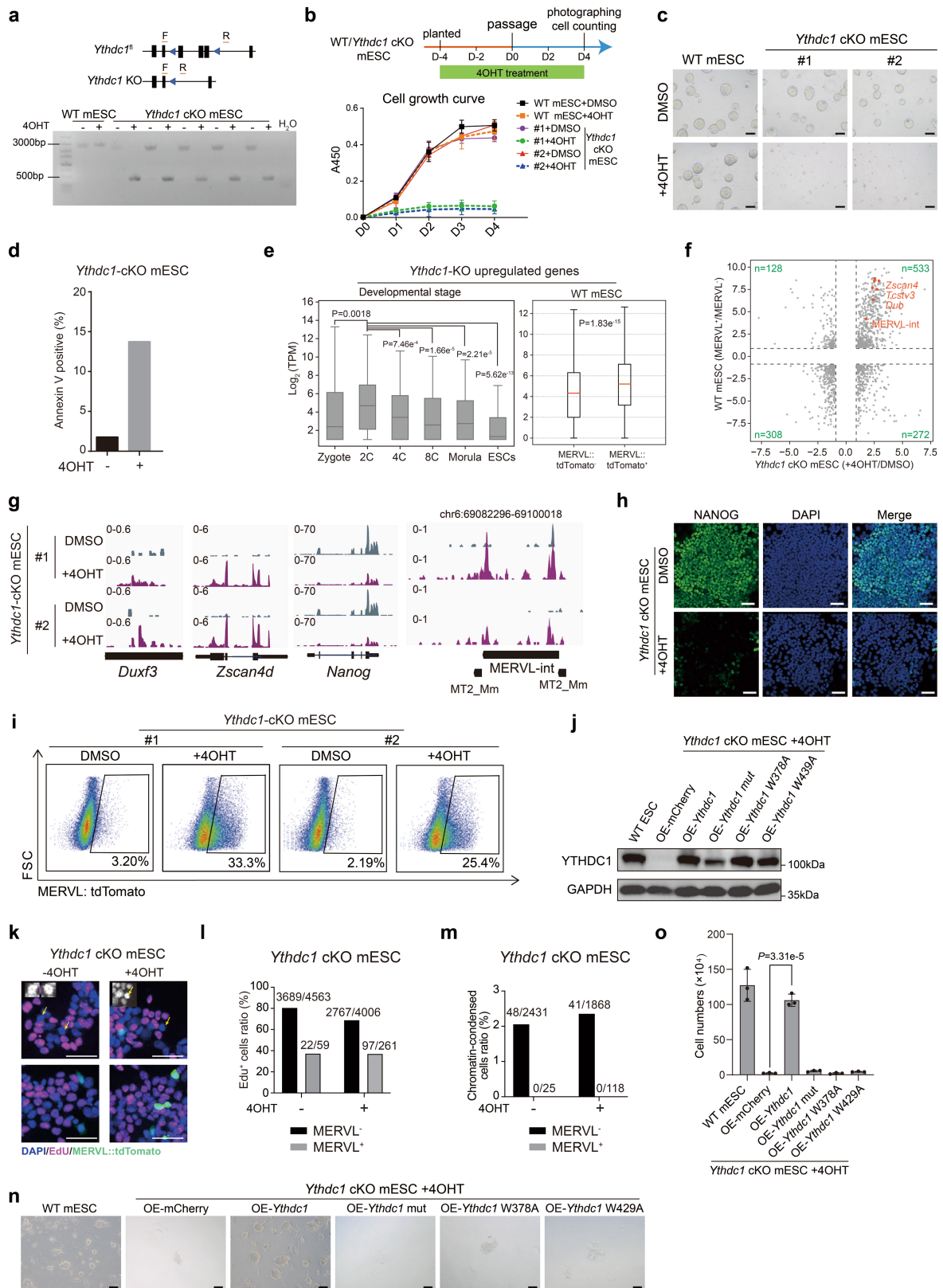
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | YTHDC1 depletion upregulates 2C-like genes and impairs proliferation of mouse ES cells. a, PCR-based genotyping assay validates the generation of the *Ythdc1*-KO cell line. Genotyping assay was repeated at least twice with similar results. **b**, Upper: schematic diagram depicts the timeline of DMSO or 4OHT treatment. Lower: growth curves of mouse ES cells after *Ythdc1*-KO. Data are mean ± s.d. (*n* = 6 independent wells pooled from two independent experiments). **c**, Phase-contrast images of WT and *Ythdc1* cKO mouse ES cells showing the viability of cell colonies. Micrographs was repeated at least three times with similar results. Scale bar, 200 μm. **d**, The ratio of annexin V positive cells after loss of *Ythdc1* (*n* = 2 independent experiments). **e**, Expression changes of the *Ythdc1*-KO upregulated genes across indicated embryonic developmental stages (left), or between MERVL⁺/MERVL⁻ cells in WT mouse ES cells (right). *P* from one-sided Mann–Whitney U test. Boxplots denote the medians and the interquartile ranges (IQR). The whiskers of a boxplot are the lowest datum still within 1.5 IQR of the lower quartile and the highest datum still within 1.5 IQR of the upper quartile. Data were from *n* ≥ 2 independent experiments. **f**, Scatter plot for the differentiate expressed genes in *Ythdc1* cKO and MERVL⁺/MERVL⁻ WT mouse ES cells.
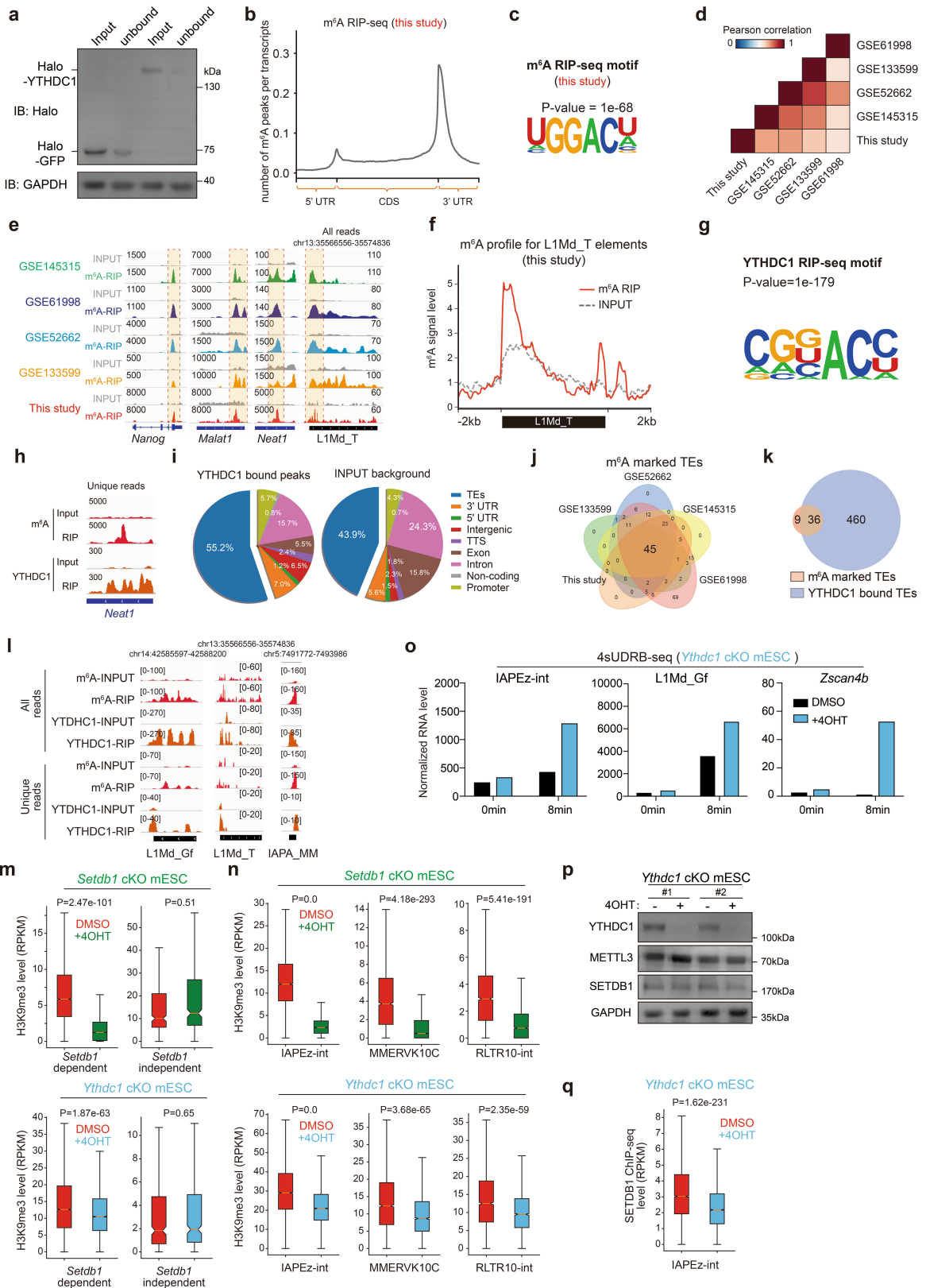
g, RNA-seq data showing the expression changes of indicated genes and MERVL-int. **h**, NANOG immunostaining in the *Ythdc1*-cKO mouse ES cells. Micrographs was repeated at least twice with similar results. Scale bar, 50 μm. **i**, Flow cytometry analysis showing the percentage of MERVL::tdTomato (*x*-axis) positive cells after loss of *Ythdc1*. **j**, Western blot validation of YTHDC1 and its mutants in the rescued assays. Immunoblotting was repeated at least twice with similar results. **k**, Fluorescence images display the cells by staining with EdU and DAPI. Micrographs were repeated at least twice with similar results. Scale bar, 50 μm. **l**, The percentage of Edu+/DAPI cells in MERVL⁺ and MERVL⁻ cells upon *Ythdc1*-KO. Number of counted cells are labelled on the top of the bar (*n* = 2 independent experiments). **m**, The percentage of phases/DAPI in WT and *Ythdc1*-KO cells. Numbers of counted cells are labelled on the top of the bar (*n* = 2 independent experiments). **n**, Phase-contrast images present the rescue effects of WT or mutant *Ythdc1* overexpression (OE) upon *Ythdc1*-KO. Micrographs were repeated at least three times with similar results. Scale bar, 200 μm. **o**, Cell viability of mouse ES cells in **n**. Data are mean ± s.d. of three independent experiments. *P* values from two-sided Student's *t*-test.

a

Morphology

Ythdc1 cKO mESCs（mCherry+）

b

P-value = 0.0048

c

Ythdc1 cKO mESCs +4OHT

mCherry CDX2 DAPI mCherry/CDX2 Merge

d

Ythdc1

Relative to *Gapdh*

e

MERVL::tdTomato (%)

P=4.58e-5

P=0.023

Ythdc1-cKO mESCs

f

Fluorescence microscope

Phase mCherry ELF5 DAPI Merge

E6.5 embryo

siNC

siNC 1#

si*Ythdc1*

si*Ythdc1* 1#

si*Ythdc1* 2#

si*Ythdc1* 3#

Confocal imaging

mCherry ELF5 DAPI Merge

siNC 1#

si*Ythdc1* 1#

g

P-value = 0.0047

% of Chimeric embryos

E6.5 embry

Trophectoderm + ICM

ICM only

siNC si*Ythdc1*

Extended Data Fig. 2 | *Ythdc1*-KO cells incorporate into trophectoderm.
a, Fluorescence images of blastocysts that have been injected with the indicated mouse ES cells at the 8-cell state. Micrographs were repeated at least twice with similar results. **b**, A column graph showing the percentages of chimeric embryos with injected ES cells incorporated into ICM or trophectoderm. Two-sided Fisher's exact test. **c**, Same as Fig. 1j, the E4.5 blastocysts developing from morula injected with *Ythdc1*-KO ES cells were stained for mCherry and CDX2. DAPI stains the nucleus. Micrographs were repeated at least twice with similar results. Scar bar, 20 μm. **d**, RT–qPCR detects the expression of *Ythdc1* upon siRNA treatment, with siNC as the negative control (*n* = 2 independent experiments). **e**, A column graph quantifies the percentages of MERVL–tdTomato

positive cells upon *Ythdc1* knockout or knockdown. The data are presented as the mean ± s.d.; measurements from *n* = 3 independent experiments. *P* values determined by two-sided Student's *t*-test. **f**, Whole-mount fluorescence imaging of representative 6.5 dpc ES cell chimeric embryos upon si*Ythdc1* or siNC. Fluorescence microscope (left) and confocal microscopy images (right) of siNC/si*Ythdc1* ES cells (mCherry⁺) in the ELF5-expressing ExE in the 6.5 dpc chimeras (ELF5 marks diploid trophoblast)[52,53]. DAPI stains the nucleus. Arrows point to mCherry⁺ cells in the ExE. Micrographs were repeated at least twice with similar results. **g**, A column graph quantifies the percentage of 6.5 dpc ES cell chimeric embryos upon si*Ythdc1* or siNC. *P* value was from two-sided Fisher's exact test.
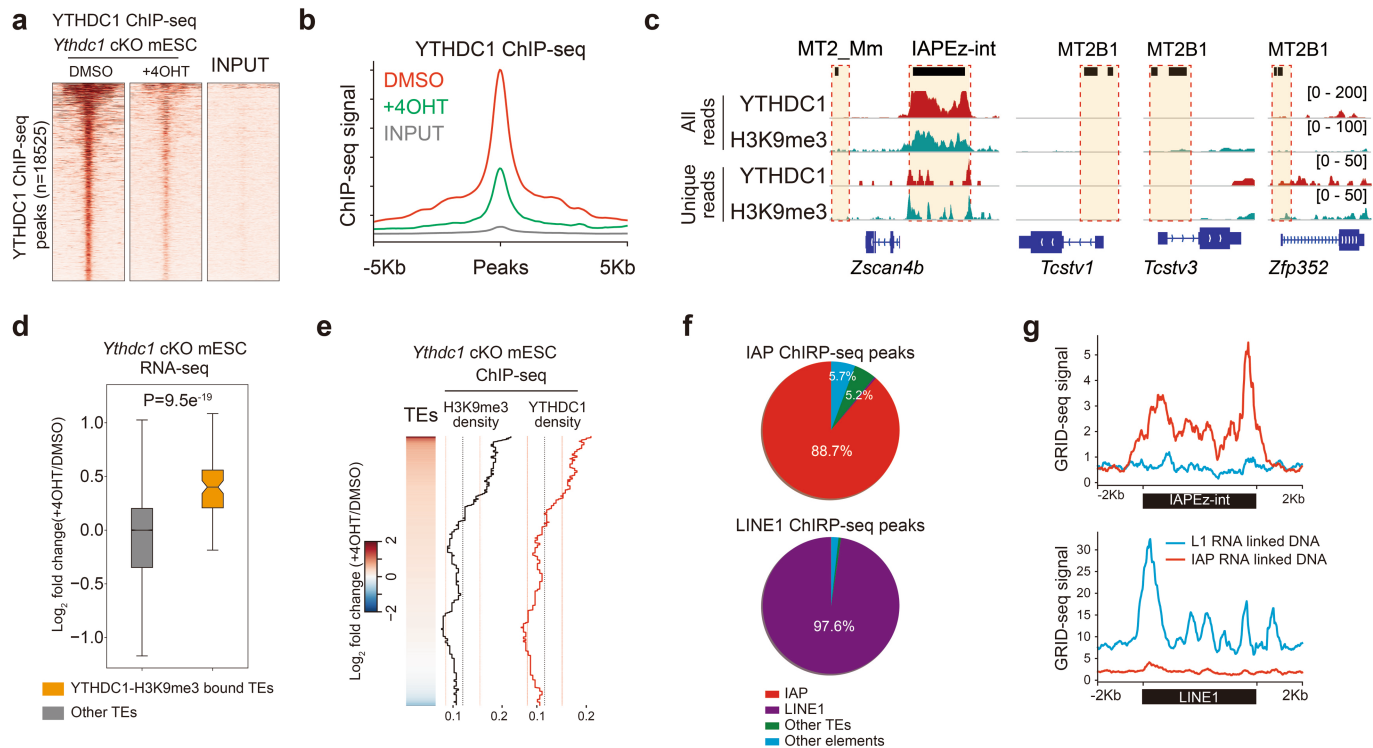
a

Halo-YTHDC1

IB: Halo

Halo-GFP

IB: GAPDH

b m⁶A RIP-seq (this study)

c **m⁶A RIP-seq motif** (this study)

P-value = 1e-68

d Pearson correlation

e

GSE145315
GSE61998
GSE52662
GSE133599
This study

Nanog   *Malat1*   *Neat1*   L1Md_T

f m⁶A profile for L1Md_T elements (this study)

g **YTHDC1 RIP-seq motif**

P-value=1e-179

h Unique reads

Neat1

i YTHDC1 bound peaks   INPUT background

j m⁶A marked TEs

k

l

L1Md_Gf   L1Md_T   IAPA_MM

o 4sUDRB-seq (*Ythdc1* cKO mESC)

IAPEz-int   L1Md_Gf   *Zscan4b*

m *Setdb1* cKO mESC

Ythdc1 cKO mESC

n *Setdb1* cKO mESC

Ythdc1 cKO mESC

p *Ythdc1* cKO mESC

YTHDC1
METTL3
SETDB1
GAPDH

q *Ythdc1* cKO mESC

P=1.62e-231

IAPEz-int

Extended Data Fig. 3 | See next page for caption.

Article

Extended Data Fig. 4 | YTHDC1 binds to H3K9me3-marked TEs chromatin.
a, Heat maps illustrating the density of YTHDC1 ChIP-seq reads upon *Ythdc1* depletion. **b**, Read count tag density pileups of YTHDC1 ChIP-seq upon *Ythdc1* depletion. **c**, Selected genomic views of YTHDC1 and H3K9me3 ChIP-seq data for the indicated TEs/genes with all mapped reads (All reads) or only unique mapped reads (Unique reads). **d**, A boxplot showing the expression of TEs with or without H3K9me3/YTHDC1 upon *Ythdc1* depletion. Statistics were determined by one-sided Mann–Whitney *U* test. Data were from *n* = 2 independent experiments **e**, Left, RNA-seq showing the change in expression of TEs upon *Ythdc1* depletion. TEs were ranked from upregulated to downregulated. Right, a moving-window average plot of the density of H3K9me3 and YTHDC1 binding. **f**, Pie charts showing the genomic distribution of the peaks identified by IAP and LINE1 ChIRP-seq. **g**, Read count tag density pileups of GRID-seq signal enriched on indicated TEs.
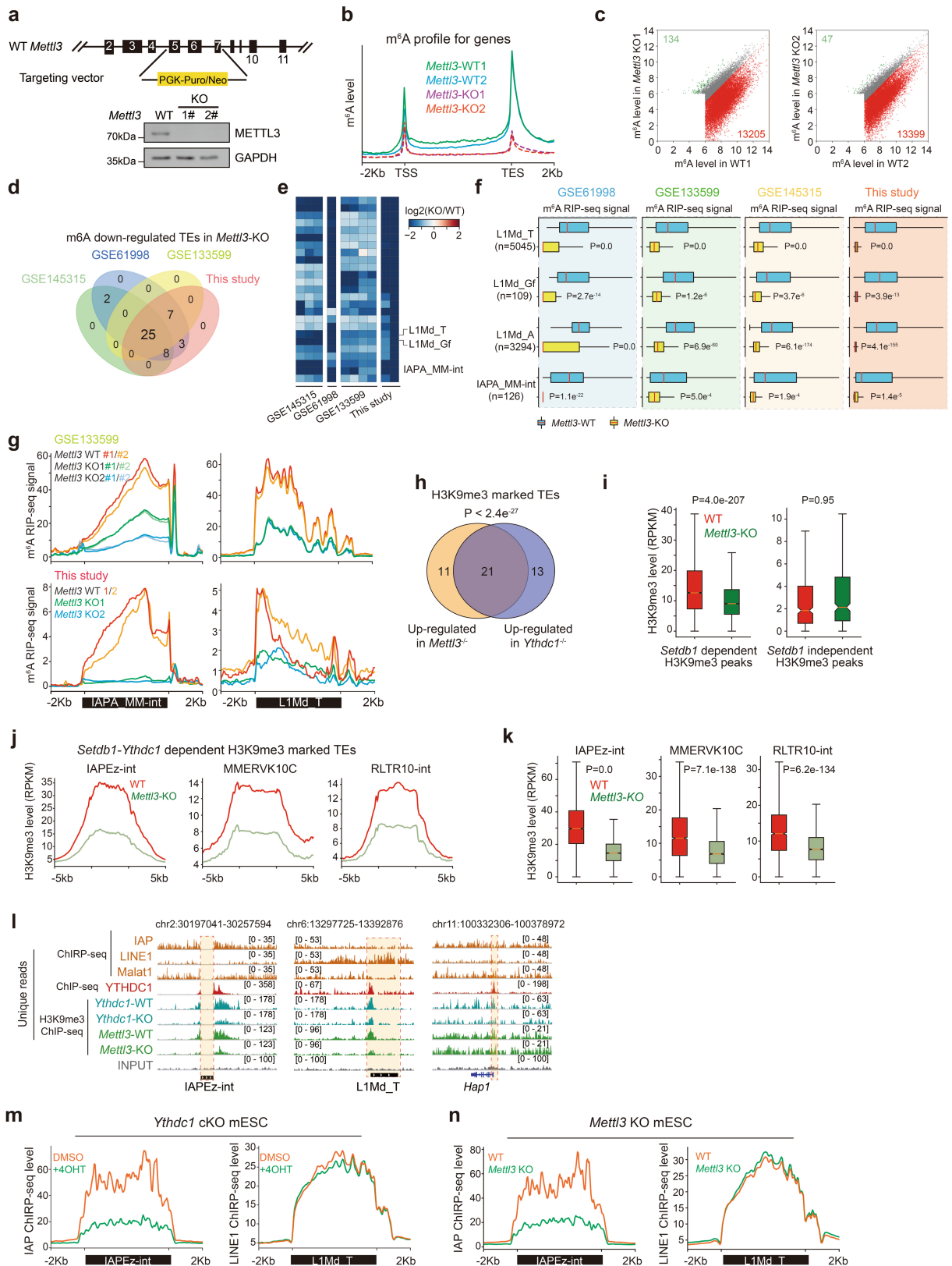
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | *Mettl3*-dependent m⁶A modification on H3K9me3-marked TEs. a, Top, a schematic diagram showing the strategy of *Mettl3* KO. Bottom, western blot validates the generation of the *Mettl3*-KO cell line. Immunoblotting was repeated at least three times with similar results. **b**, Metagene profiles of m⁶A signal along transcripts in two replicates for WT and *Mettl3* KO mouse ES cells. **c**, Scatter plots showing the m⁶A level of m⁶A peak regions in WT and *Mettl3* KO mouse ES cells. Hypermethylated peaks (green) and hypomethylated peaks (red) with m⁶A upon *Mettl3* KO are shown. **d**, The overlap of m⁶A decreased TEs after *Mettl3*-KO from 4 independent studies. **e**, The m⁶A changes for TEs upon *Mettl3* depletion from 4 independent studies. **f**, The m⁶A signal changes upon loss of *Mettl3* in different studies for indicated TEs. Boxplots denote the medians and the interquartile ranges (IQR). The whiskers of a boxplot are the lowest datum still within 1.5 IQR of the lower quartile and the highest datum still within 1.5 IQR of the upper quartile. *P* values were from two-sided Student's *t*-test. Data were from $n \geq 2$ independent experiments. **g**, Read count tag density pileups of m⁶A RIP-seq reads for indicated TEs upon *Mettl3* depletion from two independent studies. **h**, The overlap of *Ythdc1* and *Mettl3*-dependent H3K9me3 marked TEs. *P* value is from one-tailed Fisher's exact test. **i**, H3K9me3 level changes for *Setdb1*-dependent and -independent regions from Fig. 2e after *Mettl3*-KO ($n = 2$). *P* values are from two-sided Student's *t*-test. Boxplots denote the medians and the interquartile ranges (IQR). The whiskers of a boxplot are the lowest datum still within 1.5 IQR of the lower quartile and the highest datum still within 1.5 IQR of the upper quartile. **j**, Tag density pileups of H3K9me3 ChIP-seq reads for indicated TEs upon *Mettl3* depletion. **k**, Same as panel **i**, but for different TEs. **l**, Same as Fig. 3i, but only the unique mapped reads were kept. **m**, Tag density pileups of IAP (left) and LINE1 (right) ChIRP for indicated TEs upon *Ythdc1*-KO. **n**, Same as panel **m**, but for *Mettl3*-KO.
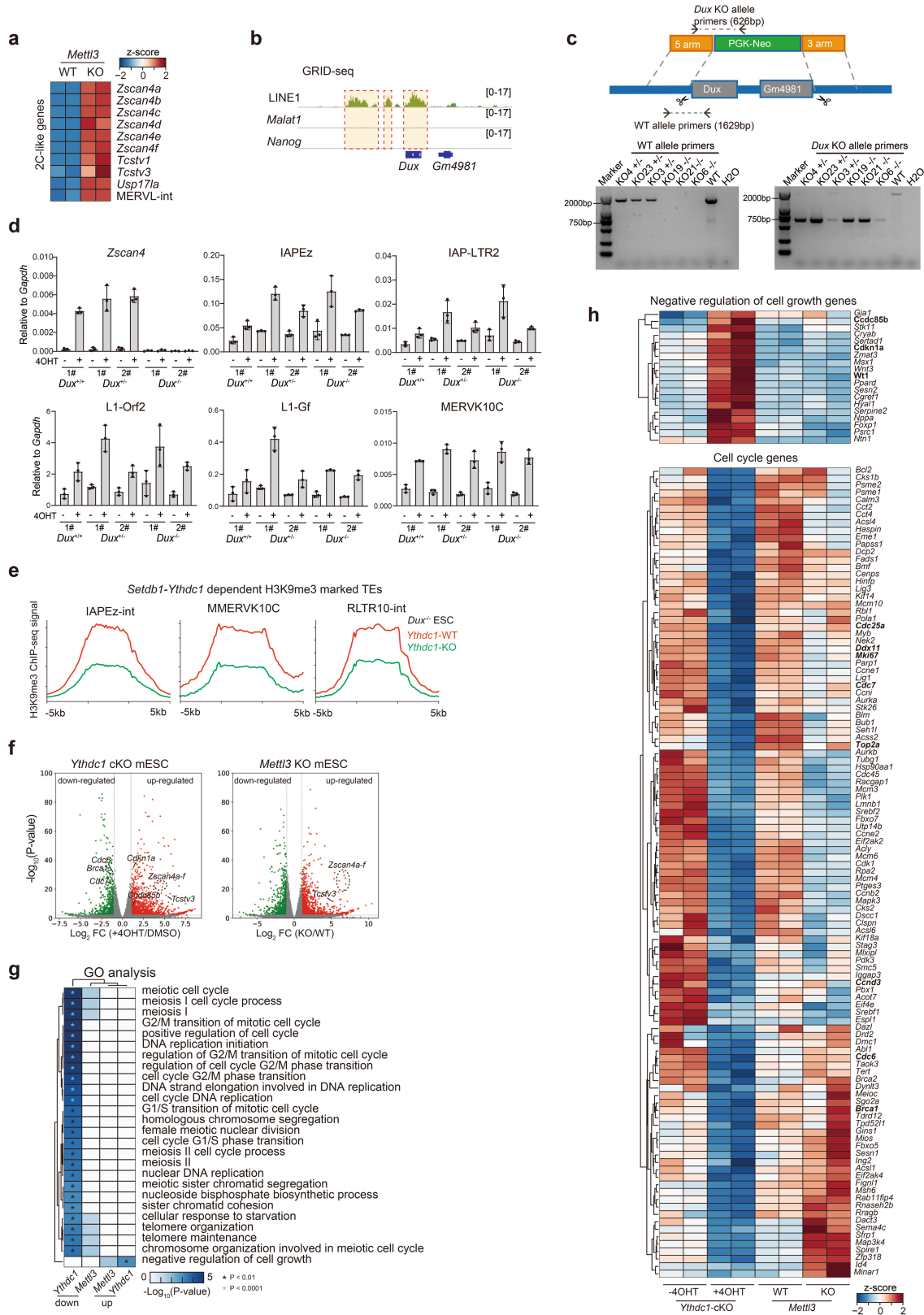
Extended Data Fig. 6 | *Dux*-KO does not impair YTHDC1–H3K9me3-mediated TE repression. a, Heat map showing the expression of indicated genes upon *Mettl3* depletion. **b**, Genome browser plot showing the GRID-seq signal in *Dux* loci. **c**, Top, a schematic diagram showing the strategy of *Dux* knockout. Bottom, PCR-based genotyping assay validates the generation of the *Dux*-KO cell line at the genome. Genotyping assay was repeated at least twice with similar results. **d**, RT–qPCR showing the expression of select genes/TEs upon

Dux depletion in *Ythdc1* cKO mouse ES cells treated with DMSO or 4OHT. Data are mean ± s.d. of three independent experiments. **e**, Read count tag density pileups of H3K9me3 ChIP-seq (from two replicates) reads for indicated TEs in *Dux* KO cells upon *Ythdc1* depletion. **f**, Volcano plot showing the differential expressed genes after loss of *Ythdc1* (left) or *Mettl3* (right). **g**, GO analysis of the genes from panel **f**. **h**, Heat map showing the expression change of indicated genes after loss of *Ythdc1* or *Mettl3*.

nature research

Corresponding author(s):	Jiekai Chen
Last updated by author(s):	Jan 24, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, seeAuthors & Referees and theEditorial Policy Checklist .

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	✘	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	✘	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	✘	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
✘	☐	A description of all covariates tested
✘	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
✘	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	✘	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
✘	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
✘	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
✘	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	RSEM(v1.2.22), bowtie2(v2.2.5), STAR(2.5.2a), DESeq2 (1.22.1), EDASeq(2.16.0), deeptools(2.5.4), HISAT2 (v2.2.0), TrimGalore (v0.6.4), MACS2 (v2.1.2), samtools (v1.9).
Data analysis	Follow the software documentations, no special changes

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The authors declare that the data and codes supporting the results of this study are available within the article and its supplementary information files or from the corresponding author upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

✘

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was determined based on previously published studies and ensure reproducibility. All experiments in this study were performed at least 2 and typically 3 independently unless otherwise specified.
Data exclusions	No data exclusions.
Replication	All attempts at replication were successful.
Randomization	This study was based on molecular and cellular biology techniques, which did not involve live organisms and no randomization techniques were required.
Blinding	The results were obtained using objective quantitative methods, so the researchers were not blinded to sample allocation during experiments and outcome assessment.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Primary antibodies Anti-Flag Sigma F1804, Anti-YTHDC1 CST 87459S, Anti-METTL3 Proteintech 15073-I-AP, Anti-SETDB1 GeneTex GTX115305, Anti-GAPDH Bioworld AP2063, Anti-Halo Promega G9211, Anti-HA CST 3724S, Anti-NANOG Bethy A300-397A, Anti-Flag CST 14793S, Anti-H3K9me3 Abcam Ab8898, Anti-YTHDC1 CST 77422, Anti-m6A NEB E1610S, Anti-ELF5 Santa-Cruz sc-166653 , Anti-mCherry LS-C204207 LSBio. Anti-CDX2 BioGenex MU392A-UC, Anti-SETDB1 Proteintech, 11231-1-AP, Anti-Mouse IgG Abcam ab18413, Anti-Rabbit IgG Abcam ab37415. Secondary antibodies Donkey anti Goat IgG 568 invitrogen A-11057, Donkey anti Mouse IgG 488 invitrogen A32766, Goat anti Rabbit IgG HRP Aksomics KC-RB-035, Goat anti Mouse IgG HRP Aksomics KC-MM-035.
Validation	All antibodies used in this study are commercially available and have been validated by manufacturer. Antibody validations and validation criteria are available on the following websites: Anti-Flag https://www.sigmaaldrich.com/catalog/product/sigma/f1804?lang=zh&region=CN Anti-YTHDC1 https://www.cellsignal.cn/products/primary-antibodies/ythdc1-p561-antibody/87459?site-search-type=Products&N=4294956287&Ntt=87459s&fromPage=plp&_requestid=2667843 Anti-METTL3 http://www.ptgcn.com/products/METTL3-Antibody-15073-1-AP.htm Anti-SETDB1 https://www.genetex.cn/Product/Detail/SETDB1-antibody-N1-N-term/GTX115305 Anti-GAPDH https://www.bioworlde.com/GAPDH-polyclonal-antibody(AP0063).html Anti-Halo https://www.promega.com.cn/Products/Protein-Detection/Primary and Secondary Antibodies/Anti-HaloTag-Monoclonal-Antibody/?catNum=G9211 Anti-HA https://www.cellsignal.cn/products/primary-antibodies/ha-tag-c29f4-rabbit-mab/3724?site-search-type=Products&N=4294956287&Ntt=3724s&fromPage=plp&_requestid=2668171 Anti-NANOG https://www.bethyl.com/product/A300-397A/Nanog+Antibody Anti-Flag https://www.cellsignal.cn/products/primary-antibodies/dykddddk-tag-d6w5b-rabbit-mab-binds-to-same-epitope-as-sigma-s-anti-flag-m2-antibody/14793?site-search-type=Products&N=4294956287&Ntt=14793s&fromPage=plp&_requestid=2668236 Anti-H3K9me3 https://www.abcam.cn/histone-h3-tri-methyl-k9-antibody-chip-grade-ab8898.html Anti-YTHDC1 https://www.cellsignal.cn/products/primary-antibodies/ythdc1-e4i9e-rabbit-mab/77422?site-search-type=Products&N=4294956287&Ntt=77422&fromPage=plp&_requestid=2668393

Anti-m6A https://international.neb.com/products/e1610-epimark-n6-methyladenosine-enrichment-kit#Product%20Information
Anti-ELF5 https://www.scbt.com/p/elf-5-antibody-g-2
Anti-mCherry https://www.lsbio.com/antibodies/mcherry-antibody-if-immunofluorescence-ihc-wb-western-ls-c204207/212463
Anti-CDX2 https://www.labome.com/product/Biogenex/MU392A-UC.html
Anti-SETDB1 http://www.ptgcn.com/products/SETDB1-Antibody-11231-1-AP.htm
Anti-Mouse IgG https://www.abcam.cn/mouse-igg2a-kappa-monoclonal-mopc-173-isotype-control-chip-grade-ab18413.html
Anti-Rabbit IgG https://www.abcam.cn/rabbit-igg-polyclonal-isotype-control-ab37415.html
Donkey anti Goat IgG 568 https://www.thermofisher.com/cn/zh/antibody/product/Donkey-anti-Goat-IgG-H-L-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-11057
Donkey anti Mouse IgG 488 https://www.thermofisher.com/cn/zh/antibody/product/Donkey-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A32766
Goat anti Rabbit IgG HRP http://www.aksomics.com/products/secondary-antibody-1.html
Goat anti Mouse IgG HRP http://www.aksomics.com/products/secondary-antibody-1.html

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	129 mouse embryonic stem cells were derived from 3.5 d.p.c inner cell mass (ICM) from female 129 mice (Stock No.: 217, Beijing Vital River Laboratory Animal Technology Co., Ltd.) HEK293T (ATCC, #CRL-1126) NIH3T3 (ATCC, #CRL-1658)
Authentication	129 mouse embryonic stem cells were authentication by Chimera experiment.
Mycoplasma contamination	Mycoplasma detection tests were conducted routinely to ensure mycoplasma-free conditions throughout this study.
Commonly misidentified lines (See ICLAC register)	Cells used are not in th ICLAC database.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Female and male CD-1 mice were used in chimeric blastocyst assay between 8 and 12 weeks of age.
Wild animals	No wild animals were used in this study.
Field-collected samples	All animal procedures were performed according to NIH guidelines. Animals were housed in individually ventilated cages containing sterile quarter-inch corncob bedding for environmental enrichment. Mice were maintained on 12 h light/dark cycle and provided with food and water in individually ventilated units in a temperature controlled room (22 - 23℃).
Ethics oversight	Animal care and experimental protocols were approved by the Guangzhou Institutes of Biomedicine and Health Ethical Committee under NO. N2020106.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE146467 Reviewer token: evuvmckervorfuh
Files in database submission	Ythdc1-cKO.DMSO.Dux-Het-H3K9me3.rep1,Ythdc1-cKO.DMSO.Dux-Het-H3K9me3.rep2,Ythdc1-cKO.4OHT.Dux-Het-H3K9me3.rep1,Ythdc1-cKO.4OHT.Dux-Het-H3K9me3.rep2,Ythdc1-cKO.DMSO.Dux-KO-H3K9me3.rep1,Ythdc1-cKO.DMSO.Dux-KO-H3K9me3.rep2,Ythdc1-cKO.4OHT.Dux-KO-H3K9me3.rep1,Ythdc1-cKO.4OHT.Dux-KO-H3K9me3.rep2,YTHDC1(Ythdc1-KO).rep1,YTHDC1(Ythdc1-KO).rep2,Mettl3.KO.H3K9me3.rep1,Mettl3.KO.H3K9me3.rep2,Mettl3.WT.H3K9me3.rep1,Mettl3.WT.H3K9me3.rep2,YTHDC1(Ythdc1-WT).rep1,YTHDC1(Ythdc1-WT).rep2,INPUT(WT), shMettl3.H3K9me3, ChIP.Input, shScramble.H3K9me3, WT.rep1-H3K9me3, WT.rep2-H3K9me3, Ythdc1-KO.rep1-H3K9me3,Ythdc1-KO.rep2-H3K9me3, YTHDC1-RIP, YTHDC1-RIP-INPUT, Ythdc1_cKO_ESC#1_+4OHT.rep1,Ythdc1_cKO_ESC#1_DMSO.rep2, Ythdc1_cKO_ESC#1_+4OHT.rep2, Ythdc1_cKO_ESC#1_DMSO.rep1, Ythdc1_cKO_ESC#1_DMSO.rep2, Ythdc1_cKO_ESC#2_+4OHT.rep1, Ythdc1_cKO_ESC#2_+4OHT.rep2, Ythdc1_cKO_ESC#2_DMSO.rep1, Ythdc1_cKO_ESC#2_DMSO.rep2，m6A-RP.Mettl3-KO1.INPUT,m6A-RP.Mettl3-KO1.IP,m6A-RP.Mettl3-KO2.INPUT,m6A-RP.Mettl3-KO2.IP,m6A-RP.Mettl3-WT.INPUT.rep2,m6A-RP.Mettl3-WT.IP.rep2,m6A-RP.Mettl3-KO.Input,m6A-RP.Mettl3-KO.IP,m6A-RP.Mettl3-WT.Input.rep1,m6A-RP.Mettl3-WT.IP.rep1 ChIRP.Mettl3-KO.IAP.Even,ChIRP.Mettl3-KO.IAP.Odd,ChIRP.Mettl3-KO.Input,ChIRP.Mettl3-KO.LacZ.Even,ChIRP.Mettl3-

KO.LacZ.Odd,ChIRP.Mettl3-KO.Malat1.Even,ChIRP.Mettl3-KO.Malat1.Odd,ChIRP.Mettl3-WT.IAP.Even,ChIRP.Mettl3-WT.IAP.Odd,ChIRP.Mettl3-WT.Input,ChIRP.Mettl3-WT.LacZ.Even,ChIRP.Mettl3-WT.LacZ.Odd,ChIRP.Mettl3-WT.Malat1.Even,ChIRP.Mettl3-WT.Malat1.Odd,ChIRP.Mettl3-KO.LINE1.1,ChIRP.Mettl3-KO.LINE1.2,ChIRP.Mettl3-WT.LINE1.1,ChIRP.Mettl3-WT.LINE1.2, ChIRP.Ythdc1-cKO.4OHTCHIRP_Input,ChIRP.Ythdc1-cKO.4OHTIAP_Even,ChIRP.Ythdc1-cKO.4OHTIAP_Odd,ChIRP.Ythdc1-cKO.4OHTLINE1_Even,ChIRP.Ythdc1-cKO.4OHTLINE1_Odd,ChIRP.Ythdc1-cKO.4OHTLacZ_Even,ChIRP.Ythdc1-cKO.4OHTLacZ_Odd,ChIRP.Ythdc1-cKO.4OHTMalat1_Even,ChIRP.Ythdc1-cKO.4OHTMalat1_Odd,ChIRP_IAP_Even,ChIRP_IAP_Odd,ChIRP_Input,ChIRP_LINE1.Even,ChIRP_LINE1_Odd,ChIRP_LacZ_Even,ChIRP_LacZ_Odd,ChIRP_Malat1.Odd,ChIRP_Malat1_Even

| Genome browser session (e.g. UCSC) | No public available genome browser session |

Methodology

| Replicates | 2 replicates for Ythdc1 knock out, 1 replicates for Mettle3 knock down, 2 replicates for Mettl3 knock out, 2 replicates for Dux +/-, 2 replicates for Dux-/- |

| Sequencing depth | Paire-ends, 20-40M raw sequencing reads for each sample |

| Antibodies | anti-H3K9me3 antibody (Abcam, #ab8898), anti-YTHDC1 antibody (CST, 77422), anti-SETDB1 antiody (Proteintech 11231-1-AP) |

| Peak calling parameters | Peaks were called by MACS2 with default parameters with just specify the genome "-g mm" |

| Data quality | H3K9me3 ChIP-seq quality was performed by comparing to previous published dataset with pearson correlation > 0.8. And we detect significant 18788 peaks. We detect 18525 peaks for peaks for YTHDC1 ChIP-seq data. Other QC filter are detailed in the method part of the manuscript |

| Software | ChIP-seq reads were aligned to the mouse mm10 genome using bowtie247 with the options '-p 20–very-sensitive–end-to-end–no-unal–no-mixed -X 2000'. The multimapped reads were kept for TE analysis, but only the best alignment is reported for those reads, if more than one equivalent best alignment was found, then one random alignment was reported. Whereas the reads mapping to mitochondrial DNA or unassigned sequences were discarded. The Bam files of biological replicates were merged using samtools (version 1.9), and peaks were called using MACS (version 2.1.2). BigWig files were generated using deeptools53 with the RPKM normalization method. |

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

| Sample preparation | MERVL::tdTomato reporter mESCs were digested with 0.25% trypsin and resuspended by PBS with 1% FBS. |

| Instrument | LSRFortessa flow cytometer (BD Biosciences) |

| Software | FlowJo V10 software |

| Cell population abundance | One hundred thousand to distinguish the tdTomato positive and negative cells. |

| Gating strategy | Using the DMSO treated cells as negative control. |

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.