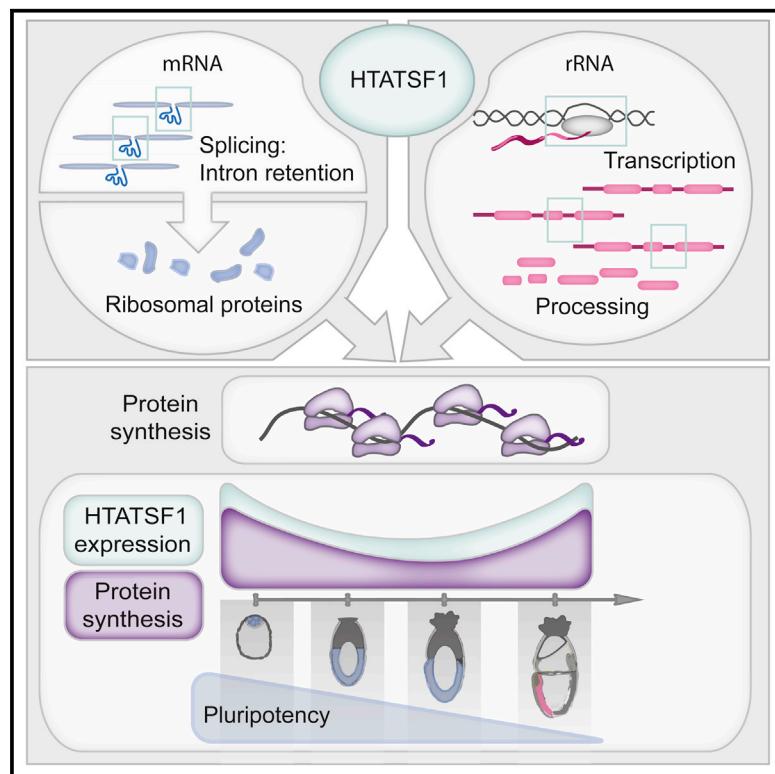


Coordinated Control of mRNA and rRNA Processing Controls Embryonic Stem Cell Pluripotency and Differentiation

Graphical Abstract



Authors

Nina S. Corsini, Angela M. Peer,
Paul Moeseneder, ...,
Hans-Christian Theussl, Isabella Moll,
Juergen A. Knoblich

Correspondence

juergen.knoblich@imba.oeaw.ac.at

In Brief

Corsini et al. identify the splicing factor HTATSF1 as a regulator of intron retention specifically in ribosomal proteins and of ribosomal RNA transcription and processing to modulate levels of overall protein synthesis. They further demonstrate that HTATSF1-mediated protein synthesis dynamics control embryonic stem cell pluripotency and neuroectoderm differentiation.

Highlights

- HTATSF1 specifically controls splicing and intron retention in ribosomal proteins
- HTATSF1 regulates ribosomal RNA transcription and processing
- HTATSF1 is required for efficient protein synthesis
- HTATSF1-dependent protein synthesis controls pluripotency and differentiation



Coordinated Control of mRNA and rRNA Processing Controls Embryonic Stem Cell Pluripotency and Differentiation

Nina S. Corsini,¹ Angela M. Peer,¹ Paul Moeseneder,¹ Mykola Roiuk,³ Thomas R. Burkard,^{1,2} Hans-Christian Theussl,^{1,2} Isabella Moll,³ and Juergen A. Knoblich^{1,4,*}

¹Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna Biocenter (VBC), Dr. Bohr-Gasse 3, 1030 Vienna, Austria

²Research Institute of Molecular Pathology (IMP), Vienna Biocenter (VBC), Campus-Vienna-Biocenter 1, 1030 Vienna, Austria

³Department of Microbiology, Immunobiology and Genetics, Max F. Perutz Laboratories, University of Vienna, Vienna Biocenter (VBC), Dr. Bohr-Gasse 9, 1030 Vienna, Austria

⁴Lead Contact

*Correspondence: juergen.knoblich@imba.oewa.ac.at

<https://doi.org/10.1016/j.stem.2018.03.002>

SUMMARY

Stem cell-specific transcriptional networks are well known to control pluripotency, but constitutive cellular processes such as mRNA splicing and protein synthesis can add complex layers of regulation with poorly understood effects on cell-fate decisions. Here, we show that the RNA binding protein HTATSF1 controls embryonic stem cell differentiation by regulating multiple aspects of RNA processing during ribosome biogenesis. HTATSF1, in a complex with splicing factor SF3B1, controls intron removal from ribosomal protein transcripts and regulates ribosomal RNA transcription and processing, thereby controlling 60S ribosomal abundance and protein synthesis. HTATSF1-dependent protein synthesis is essential for naive pre-implantation epiblast to transition into post-implantation epiblast, a stage with transiently low protein synthesis, and further differentiation toward neuroectoderm. Together, these results identify coordinated regulation of ribosomal RNA and protein synthesis by HTATSF1 and show that this essential mechanism controls protein synthesis during early mammalian embryogenesis.

INTRODUCTION

Cell-fate transitions are regulated by signaling pathways that integrate extracellular signals with cell-intrinsic transcriptional programs. Cell-essential events like mRNA splicing and protein synthesis, in contrast, have been regarded as housekeeping processes common to all cells without major fate determining roles. More recently, however, unexpected roles for these processes are beginning to emerge. Here, we show that changes in protein synthesis and RNA processing can influence cell fate during early mammalian development.

Early mammalian development is characterized by a gradual loss of pluripotency. The pre-implantation epiblast produces all

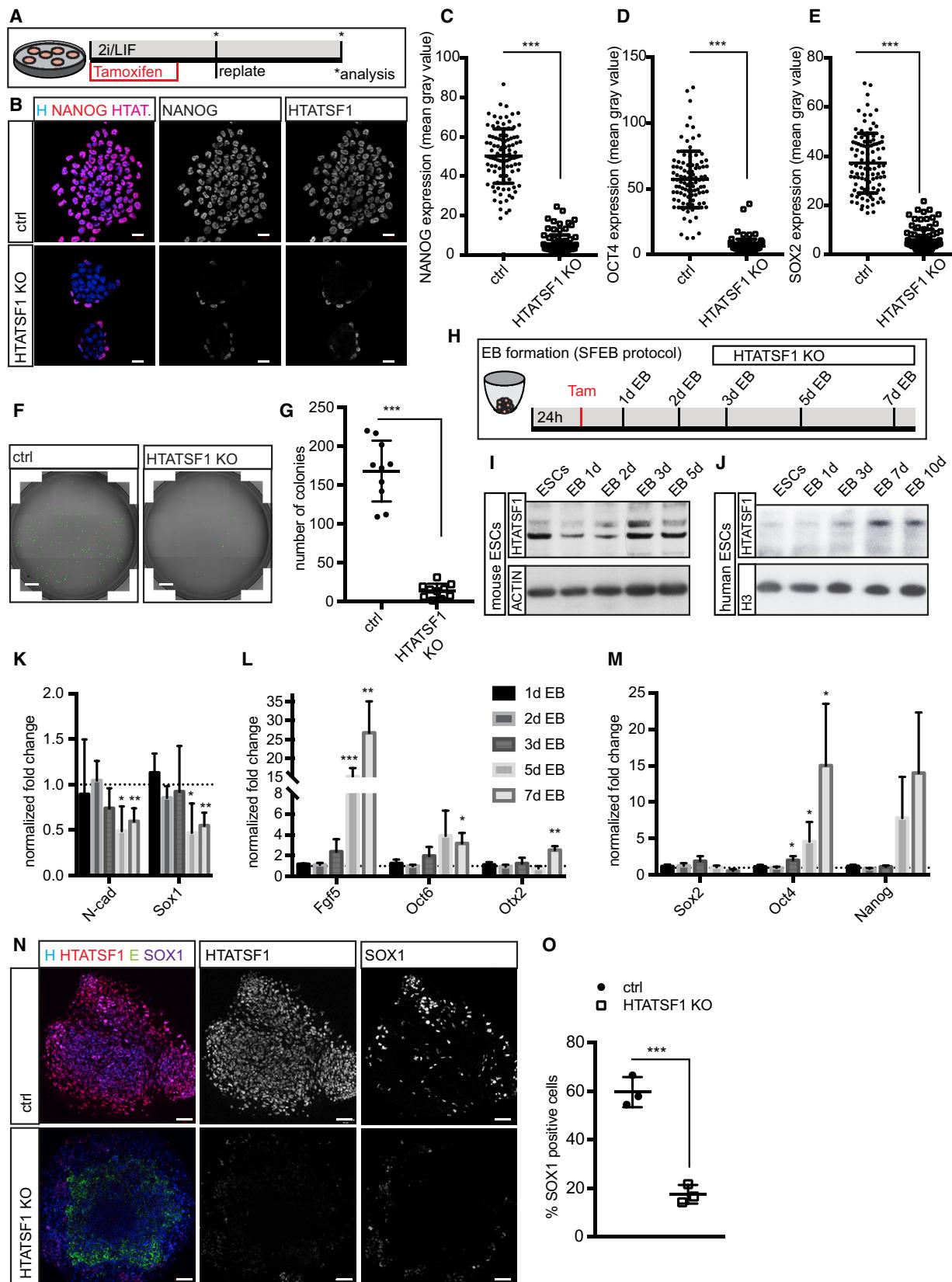
somatic and germline cells of the embryo and exemplifies a state termed “naive” pluripotency reflected by expression of a specific set of transcription factors and a hypomethylated genome (Boroviak et al., 2014; Nichols and Smith, 2012). *In vitro*, mouse embryonic stem cells (mESCs) cultured in 2i/LIF most closely resemble the naive pre-implantation epiblast (E4.5 in mouse) (Boroviak et al., 2015). After implantation, the early post-implantation epiblast undergoes major cellular changes entering a phase of “formative” pluripotency, probably best reflected *in vitro* by “epiblast-like cells” (EpiLCs) that are transcriptionally similar to the pre-gastrulation E5.5–E6.0 epiblast (Hayashi et al., 2011; Kalkan et al., 2017; Smith, 2017). Subsequently, epiblast cells start to regionally express specification factors and initiate gastrulation. This “primed” phase of late pluripotency has been studied using “epiblast stem cells” (EpiSCs), a heterogeneous population of cells with transcriptional similarities to the late gastrula-stage epiblast (Kojima et al., 2014).

Conventionally derived human ESCs (hESCs) are in a primed pluripotent state *in vitro* but can be converted to naive pluripotency (Takashima et al., 2014; Weinberger et al., 2016). Both, human and mouse ESCs can be differentiated into the three germ-layers ectoderm, mesoderm, and endoderm and thus provide an accessible *in vitro* system to study the major regulatory processes during gastrulation.

Splicing is an mRNA processing event essential for the synthesis of most proteins. Intron retention is a specific form of alternative splicing (Braunschweig et al., 2014). Failure to remove one or more introns within a given transcript can result in mRNA decay and reduced protein production. It is currently estimated that 5%–10% of all mRNAs contain retained introns (Boutz et al., 2015). Intron retention allows the rapid adjustment of transcript and protein levels without the need for transcriptional changes. It is involved in diverse biological functions ranging from granulocyte differentiation to neuronal excitability control (Bell et al., 2010; Braunschweig et al., 2014; Mauger et al., 2016; Wong et al., 2013). The mechanisms that regulate intron retention during cell-fate specification are currently not well understood.

Another basic cellular process relevant for cell-fate specification is protein synthesis. Tight control of protein translation is imperative for differentiation, proliferation, and growth of diverse types of stem cells in several organisms (Blanco et al., 2016;





(legend on next page)

Sanchez et al., 2016; Zhang et al., 2014). In hematopoietic stem cells, for example, both an increase or decrease in protein synthesis impairs proper function (Signer et al., 2014). Further, regulation of protein synthesis and protein homeostasis can orchestrate developmental changes rather than simply responding to them (Noormohammadi et al., 2018).

Our data establish the RNA binding protein HTATSF1 as a regulator of ribosomal protein levels and ribosomal RNA resulting in changes in protein synthesis. HTATSF1 is conserved from yeast to humans and contains two RNA binding RNA recognition motif (RRM) domains and an acidic C terminus. In yeast, the HTATSF1 homolog Cus2 regulates pre-spliceosome assembly (Chattoth et al., 2014). Our data show that HTATSF1 binds to the U2 small nuclear ribonucleoprotein (snRNP) complex and controls mRNA splicing of ribosomal protein genes. It also influences the production and processing of ribosomal RNA components. This results in changes of levels of the 60S large ribosomal subunit and protein synthesis. HTATSF1 and protein synthesis levels are reduced when ESCs undergo differentiation. Our data reveal a mechanism whereby constitutive, cell-essential processes can cooperate to regulate cell-fate transitions.

RESULTS

HTATSF1 Is Required for Pluripotency Maintenance

In *Drosophila*, HTATSF1/Barricade regulates stem cell lineages (Neumüller et al., 2011; Abramczuk et al., 2017). To test whether this role is evolutionarily conserved, we used mouse and hESCs as an *in vitro* system for self-renewal and differentiation. To assess whether HTATSF1 had a role in maintenance of ESC pluripotency, we asked whether HTATSF1 was expressed in mESCs. ESC cultured under 2i/LIF plus serum conditions expressed high levels of HTATSF1 (Figures 1B, 1I, and S1K). We next asked whether loss of HTATSF1 would interfere with ESC pluripotency. To delete HTATSF1 in ESCs, we generated mice bearing a conditional *Htatsf1* loss of function allele by inserting loxP sites flanking Exon3 to Exon6 (Figures S1A and S1B). We derived ESC lines from homozygous mice that carried a heterozygous constitutively expressed CreERT2 insertion (*Htatsf1*^{f/f}; *CreERT2*^{+/−}).

Upon addition of 4-OH-tamoxifen to ESCs, HTATSF1 was efficiently deleted (Figure S1C). We assessed pluripotency in ESCs by performing immunofluorescence for the pluripotency markers NANOG, OCT4, and SOX2 and testing for colony-forming efficiency (Figures 1A–1G, S1D, and S1E). While control ESCs expressed high levels of the pluripotency markers NANOG (Figures 1B and 1C), OCT4 (Figures 1D and S1D), and SOX2 (Figures 1E and S1E), these markers were reduced upon loss of HTATSF1, indicating that HTATSF1 is required for pluripotency maintenance. We next assessed colony-forming efficiency (Figures 1F and 1G). Cells lacking HTATSF1 showed a reduced ability to re-form colonies, an indication that HTATSF1 is required for maintaining pluripotent ESCs. Cells lacking HTATSF1 also accumulated in G2/M phase and after prolonged culture underwent cell death (Figure S1F, data available on request).

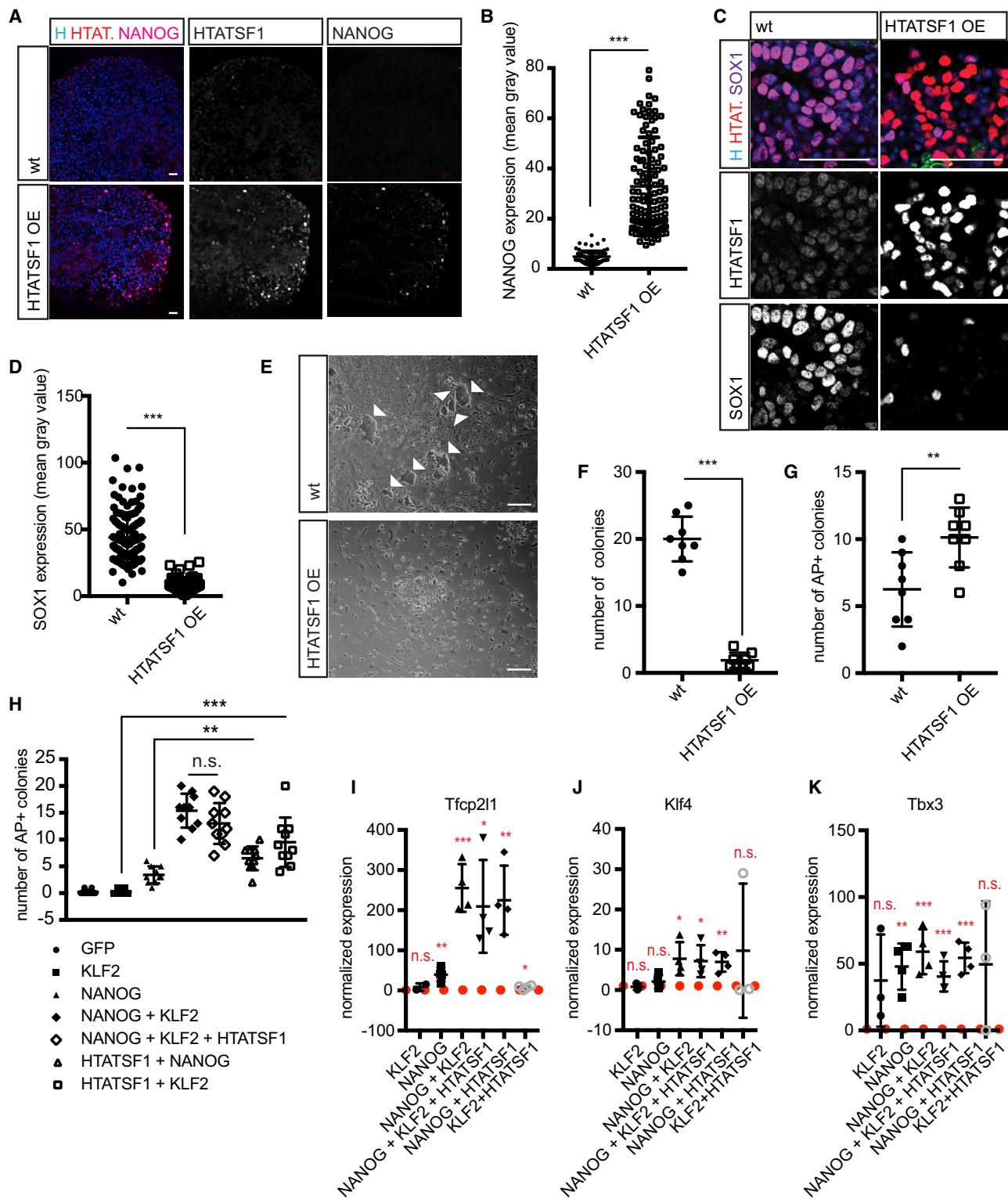
To analyze whether the observed effect on cell cycle was mediated by p53, we stained for p53 in ESCs after 4-OH-tamoxifen treatment. In contrast to control cells, HTATSF1 knockout (KO) cells showed p53 nuclear accumulation, a hallmark of p53 activation (Figure S1G). Thus, p53 signaling could potentially mediate the observed effects upon loss of HTATSF1. p53 has a role in mESC maintenance and differentiation (Lin et al., 2004). To test whether HTATSF1 acts through p53, we generated *p53*^{−/−} ESCs (*Htatsf1*^{f/f}; *CreERT2*^{+/−}; *p53*^{−/−}). Deletion of HTATSF1 in *p53*-deficient pluripotent ESCs partly rescued the cell-cycle phenotype (Figure S2B). However, levels of NANOG, OCT4, and SOX2 (Figures S2A and S2C–S2H) as well as colony-forming efficiency (Figures S2I and S2J) were still reduced upon loss of HTATSF1. These results indicate that p53 signaling was not responsible for the loss of pluripotency observed upon HTATSF1 depletion.

HTATSF1 Shows a Dynamic Expression Pattern Reflecting Developmental Potential

To understand whether HTATSF1 had a role also in ESC differentiation, we studied HTATSF1 function in embryoid bodies (EBs). EBs were cultured under the defined SFEB (serum-free culture of embryoid-body-like aggregates) conditions, which

Figure 1. HTATSF1 Is Required for Pluripotency Maintenance and Neuroectoderm Differentiation

- (A) Experimental workflow to assess pluripotency maintenance.
- (B–E) *Htatsf1*^{f/f}; *CreERT2*^{+/−} mESCs were treated with 4-OH-tamoxifen (HTATSF1 KO) or ethanol (ctrl) and analyzed 72 hr later.
- (B) Immunofluorescence for HTATSF1 and NANOG. Scale bars represent 20 μm. H, Hoechst; HTAT., HTATSF1.
- (C) Quantification of NANOG expression as shown in (B).
- (D) Quantification of OCT4 expression as shown in Figure S1D.
- (E) Quantification of SOX2 expression as shown in Figure S1E.
- (C–E) Values are shown as mean ± SD; n = 100 cells from 3 independent experiments.
- (F and G) Colony-forming efficiency assay. *Htatsf1*^{f/f}; *CreERT2*^{+/−} mESCs were treated with 4-OH-tamoxifen (HTATSF1 KO) or ethanol (ctrl) and analyzed 72 hr later. Cells were split and re-plated, and colony-forming efficiency was analyzed.
- (F) Colonies from ctrl and HTATSF1 KO cells. Scale bars represent 2 mm.
- (G) Quantification of colony-forming efficiency. Values are shown as mean ± SD; n = 10 from 4 independent experiments.
- (H) Experimental workflow to assess loss of HTATSF1 function during differentiation. *Htatsf1*^{f/f}; *CreERT2*^{+/−} mESCs were cultured as EBs. 24 hr after initial setup of EBs, cultures were treated with 4-OH-tamoxifen or ethanol and analyzed at the indicated time points.
- (I) Western blot for HTATSF1 in mESCs and EBs during differentiation. ACTIN was used as loading control.
- (J) hESCs cultured under pluripotency conditions or differentiated as EBs for the indicated time points. Histone H3 is used as loading control.
- (K–M) qRT-PCR for N-cad and Sox (K), Fgf5, Oct6, and Otx2 (L), Sox2, Oct4, and Nanog (M). Data are normalized to housekeeping gene Srp14 and normalized to control (ethanol treatment) (ctrl = 1; represented as dashed line). Values are shown as mean ± SD; n = 3 independent experiments with 96 EBs each (except Oct4 and Sox1 [1–5 day EB]: n = 4). N-cad, N-Cadherin.
- (N) Immunohistochemistry for HTATSF1, E-CADHERIN, and SOX1 7 days after treatment. Scale bars represent 50 μm. H, Hoechst; E, E-CADHERIN.
- (O) Quantification of SOX1 positive cells as shown in (N). Values are shown as mean ± SD; n = 3.

**Figure 2. HTATSF1 in Naïve and Primed Pluripotency**

(A–D) Wild-type (wt) and ESCs overexpressing HTATSF1 (HTATSF1 OE) were cultured as EBs and analyzed 8 days after initial EB setup.

(A) Immunohistochemistry for NANOG and HTATSF1. Scale bars represent 20 μ m. H, Hoechst; HTAT., HTATSF1.(B) Quantification of NANOG immunohistochemistry shown in (A). Values are shown as mean \pm SD; n = 100 cells from 3 independent experiments.(C) Immunohistochemistry for SOX1 and HTATSF1. Scale bars represent 50 μ m. H, Hoechst; HTAT., HTATSF1.(D) Quantification of SOX1 immunohistochemistry shown in (C). Values are shown as mean \pm SD; n = 100 cells from 3 independent experiments.

(legend continued on next page)

allows generation of predominantly neuroectoderm tissue after 5–7 days of culture (Kamiya et al., 2011). During the course of EB differentiation, EBs lose markers of the pluripotent pre-implantation epiblast (Nanog) and start expressing markers of the post-implantation epiblast (Otx2, Oct6, Fgf5) before giving rise to neuroectoderm (N-cadherin) (Figure S1M). In mESCs cultured under pluripotency conditions, HTATSF1 was highly expressed (Figures 1I and S1K). Upon differentiation, HTATSF1 expression decreased transiently with lowest expression observed at 1 and 2 days EB differentiation (corresponding to 2 and 3 days after initiation of EB culture, see Figure 1H), a stage corresponding to the post-implantation epiblast (Figures 1I and S1K). When the cultures progressed further, HTATSF1 levels increased (Figures 1I and S1K). Thus, HTATSF1 shows a dynamic expression pattern with two peaks during naive pluripotency and neuroectoderm generation and low expression in the post-implantation epiblast. In hESCs, in contrast, HTATSF1 levels were low under pluripotency conditions and increased during neuroectoderm differentiation (Figure 1J), consistent with the fact that hESCs are similar to mouse post-implantation epiblast stem cells. We further asked whether loss of HTATSF1 was a general phenomenon during differentiation. We assessed expression of HTATSF1 and NANOG upon withdrawal of 2i/LIF. NANOG levels decreased 24 hr after 2i/LIF withdrawal (Figures S1H and S1J). Levels of HTATSF1 started to decrease at 24 hr and were low at 48 hr after 2i/LIF withdrawal (Figures S1H and S1I), indicating that loss of HTATSF1 is indeed a more general phenomenon observed upon exit from naive pluripotency.

HTATSF1 Is Required for Neuroectoderm Differentiation

Our results showed that HTATSF1 is required for maintaining pluripotency in mESCs. HTATSF1 expression peaked not only during pluripotency, but also during neuroectoderm development. We therefore set out to prevent re-expression of HTATSF1 by deleting HTATSF1 at the stage that corresponds to the post-implantation epiblast characterized by low HTATSF1 expression (Figures 1I and S1K). We treated EBs with 4-OH-tamoxifen 24 hr after inducing EB formation, which results in HTATSF1 deletion from the EB 2-day time point, thus preventing re-expression of HTATSF1 (Figures 1H and S1L). Loss of HTATSF1 resulted in reduced expression of the neuroectoderm markers N-cadherin and Sox1 at the 5- and 7-day EB time point (Figure 1K). We also observed a failure to downregulate the pluripotency marker Oct4 upon loss of HTATSF1, while Sox2 levels were unchanged (Figure 1M). We further found that loss of HTATSF1 resulted in failure to downregulate formative pluripotency markers Otx2 and Oct6 as well as continued expression of the epiblast marker Fgf5 (Figure 1L). Taken together, loss of HTATSF1 results in a reduction of neuroectoderm marker expression and failure to downregulate markers of pluripotency, formative pluripotency, and epiblast, indicating that HTATSF1-deficient cells fail to effi-

ciently differentiate toward neuroectoderm. To confirm these findings, we performed immunohistochemistry for the neuroectoderm marker SOX1. While immunohistochemistry of control EBs revealed an abundance of SOX1 positive cells, HTATSF1 KO EBs were mostly negative for SOX1 (Figures 1N and 1O), confirming that HTATSF1 is required for efficient neuroectoderm generation. We did not find differences in the generation of mesoderm and endoderm upon loss of HTATSF1 (Figures S1N–S1Q), indicating that HTATSF1 is not involved in these processes. We further asked whether failure to generate neuroectoderm was due to p53 induction. Deletion of HTATSF1 in *Htatsf1*^{f/f}; *CreERT2*^{+/-}; *p53*^{-/-} cells resulted in a failure to generate SOX1 positive neuroectoderm (Figures S2K–S2M), suggesting that p53 was not involved in the failure to generate neuroectoderm upon HTATSF1 loss.

HTATSF1 Overexpression Prevents Differentiation

Our data show that differentiation of mESCs is characterized by a transient stage of low expression of HTATSF1. To test whether the decrease in HTATSF1 expression was a prerequisite for differentiation, we generated ESC lines stably overexpressing HTATSF1 (HTATSF1 OE). We analyzed wild-type (WT) and HTATSF1 OE EBs at the 7-day EB time point (Figures 1H and 2A–2D). While WT cells had lost expression of NANOG, indicating differentiation, cells overexpressing HTATSF1 retained NANOG expression, indicating failure to differentiate (Figures 2A and 2B). We further assessed neuroectoderm generation. WT EBs were positive for the neuroectoderm marker SOX1, while ESCs overexpressing HTATSF1 failed to induce SOX1, confirming that HTATSF1 OE cells did not differentiate into neuroectoderm (Figures 2C and 2D). We next tested differentiation toward endoderm and mesoderm. While WT EBs were positive for the endoderm marker SOX17 and the mesoderm marker BRACHYURY (T), cells overexpressing HTATSF1 failed to induce SOX17 or T, suggesting that overexpression of HTATSF1 prevented differentiation toward all three germ layers (Figures S3A–S3D). Thus, the transient downregulation of HTATSF1 observed upon exit from pluripotency is required for efficient mESCs differentiation.

HTATSF1 in Naive and Primed Pluripotency

Our data suggest a role for HTATSF1 during naive mESC pluripotency and differentiation. We next asked whether HTATSF1 was involved in the establishment of primed EpiSCs from naive ESCs. EpiSCs expressed lower levels of HTATSF1 than naive ESCs (Figure S3E) and were characterized by high Fgf5 expression (Figure S3F). We tested the effects of forced HTATSF1 expression on the generation of EpiSCs from ESCs. While WT cells generated EpiSC colonies, HTATSF1 overexpression prevented generation of EpiSCs (Figures 2E and 2F), indicating that downregulation of HTATSF1 is crucial for the establishment

(E) EpiSC generation from WT and HTATSF1-overexpressing mESCs. EpiSC colonies are identified by flat morphology (arrowheads). Scale bar represents 500 μm.

(F) Quantification of EpiSC generation. Values are shown as mean ± SD; n = 8.

(G) WT and HTATSF1-overexpressing EpiSC cells were reverted to rESCs. Values are shown as mean ± SD; n = 8.

(H) Generation of human naive ESCs from human ESCs. Values are shown as mean ± SD; n = 10.

(I–K) qRT-PCR for Tfcp2l1 (I), Klf4 (J), and Tbx3 (K). Data are normalized to Actin and to ctrl (H9 cells grown under standard ESC conditions) (ctrl = 1; represented as red dashed line). Values are shown as mean ± SD; n = 4 independent experiments [except for KLF2 n = 3].

of EpiSCs. Conversely, HTATSF1 overexpression in EpiSCs facilitated their conversion into reverted ESCs (rESCs) (Figure 2G). We next asked whether overexpression of HTATSF1 could enhance generation of human naive pluripotent stem cells. In hESCs, which express low levels of HTATSF1 (Figure 1J), expression of HTATSF1 together with either NANOG or KLF2 allowed the formation of AP positive colonies when cultured in 2i/LIF+Gö (Takashima et al., 2014) (Figure 2H). To assess whether AP positive cells exhibited characteristics of naive pluripotent stem cells, we analyzed expression of naive pluripotency markers Tfcp2l1, Klf4, and Tbx3 (Figures 2I–2K). Cells expressing NANOG together with KLF2 expressed higher levels of naive pluripotency markers Tfcp2l1, Klf4, and Tbx3 than control H9 cells cultured under standard conditions. Levels of pluripotency markers Rex1, Nanog, and Oct4, which are not specific for naive pluripotency were only slightly changed (with the exception of Nanog in the NANOG overexpression conditions) (Figures S3G–S3I). Overexpression of NANOG resulted in slightly increased levels of Tfcp2l1 and increased levels of hTbx3 (Figures 2I and 2K). Although co-expression of HTATSF1 with NANOG resulted in only a slight increase of AP⁺ colonies, we observed increased expression of the three markers for naive pluripotency Tfcp2l1, Klf4, and Tbx3 (Figures 2I–2K). These results indicate that HTATSF1 promotes a naive pluripotency like state in combination with NANOG; however, the effect on the efficiency of naive pluripotent stem cells generated (Figure 2H) is less pronounced compared to overexpression of the combination of KLF2 and NANOG. Taken together, HTATSF1 has a role in the generation of naive and primed pluripotency states in both human and mouse ESCs.

HTATSF1 Controls Intron Retention in Ribosomal Protein Genes

Our results indicate that HTATSF1 is a regulator of ESC pluripotency and differentiation. To explore how HTATSF1 exerts these functions on a molecular level, we searched for its binding partners in mESCs. Using immunoprecipitation (IP) and mass spectrometry, we identified several splicing factors and proteins involved in ribosomal RNA transcription and processing as HTATSF1 interactors (Figure 3A; Table S1). HTATSF1 interacted with splicing factor SF3B1, a component of the U2 snRNP complex and with U2 snRNP complex member SF3B2 (Figures 3A and 3B; Table S1). Further, RNA IP followed by qRT-PCR analysis revealed binding of HTATSF1 to U2 snRNA (Figure 3C). Thus, our data confirm that HTATSF1 acts in the U2 snRNP complex (Fong and Zhou, 2001).

To identify potential roles for HTATSF1 in mRNA splicing, we performed RNA sequencing in control and HTATSF1 KO mESCs (Figure 3D; Table S2). We analyzed transcriptomes using two different algorithms to detect splicing defects. The analysis failed to find any evidence for alternative splice site usage or for exon skipping in HTATSF1 KO cells. However, both methods of analysis showed an increase of transcripts with retained introns at two time points (48 and 72 hr) after 4-OH-tamoxifen addition. Combining the overlap of both methods and manual confirmation led to the identification of 45 genes whose transcripts showed a significant increase in the retention of one or several introns (Table S2). We confirmed retained introns in *Rpl36a*, *Eif1*, *Eif3c*, *Rpl9*, *Rps20*, *Rps18*, and *Rpl7a* by qRT-PCR and

RT-PCR (Figures 3E, S4C, and S4D). Thus, our data suggest that HTATSF1 regulates the retention of a specific set of introns in ESCs.

To test whether the gene set affected by HTATSF1 loss correlates with a specific biological function, we used gene ontology (GO) term and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. Surprisingly, this analysis revealed a striking specificity for proteins involved in the regulation of protein translation. Proteins identified had known functions in ribosome biogenesis and assembly, including ribosomal proteins and translation initiation and elongation factors (Figure 3F; Table S2). We next analyzed sequence features of the introns affected by HTATSF1 loss. Retained introns were more GC rich (Figure 3G) and shorter (Figure 3H) than the genome-wide average and had a strong tendency to be the first intron within the transcript (39%) (Figure S4E). We also explored the consequences of HTATSF1-regulated intron retention. Sequence analysis of transcripts with retained introns revealed that most affected introns were in the coding region, although some were in the 5' UTR (Figure S4F). For all introns in the coding region, intron retention resulted in the generation of multiple STOP codons (Table S2). Accordingly, the protein levels of RPL36A, RPL7A, EIF3C, and EIF1 were significantly reduced in HTATSF1 KO ESCs (Figures 3I and 3J), indicating that retention of introns results in reduction of protein levels. Thus, HTATSF1 functions with the U2 snRNP complex to specifically mediate the removal of introns in ribosomal proteins and translation initiation factors to control their protein levels.

HTATSF1 was described as a factor that stimulates general transcription elongation. Mass spectrometry revealed an association of HTATSF1 with RNA polymerase II (Pol II) subunit RPB1 (Table S1). Chromatin IP sequencing (ChIP-seq) revealed only minor and few differences in Pol II traveling rates between control and HTATSF1 KO cells, indicating that stimulating transcriptional elongation might not be the major function of HTATSF1 in ESCs (Figures S4G and S4H). Since HTATSF1 interacted with SF3B1, a factor whose loss can cause intron retention phenotypes, we asked whether the intron retention observed upon loss of HTATSF1 was due to a reduction in SF3B1 levels (Figures S4A and S4B). We did not observe obvious changes in SF3B1 levels upon HTATSF1 KO, indicating that the intron retention phenotype was not mediated by reduced levels of SF3B1.

HTATSF1 Affects rRNA Transcription and Processing

In addition to the U2 snRNP complex, HTATSF1 interacted with proteins involved in rRNA transcription and processing (Figure 3A; Table S1). HTATSF1 was associated with the rRNA transcriptional regulators MYBBP1A (Hochstatter et al., 2012) and DDX21 (Calo et al., 2014), as well as nucleolar proteins UTP20, FIBRILLARIN, and large subunit biogenesis factors BOP1 and PES1. HTATSF1 was detected both in nuclear and nucleolar fractions in crosslinked samples (Figure S5A). We confirmed the interaction with large subunit biogenesis factors BOP1 and PES1 by Co-IP of crosslinked nuclear extracts (Figure S5B). We next asked whether HTATSF1 associated with rRNA transcripts. We performed RNA-IP from crosslinked nucleoli (Figure 4B). We found that HTATSF1 bound to the ITS2 site. Removal of the ITS2 is a prerequisite for the maturation of the 28S rRNA

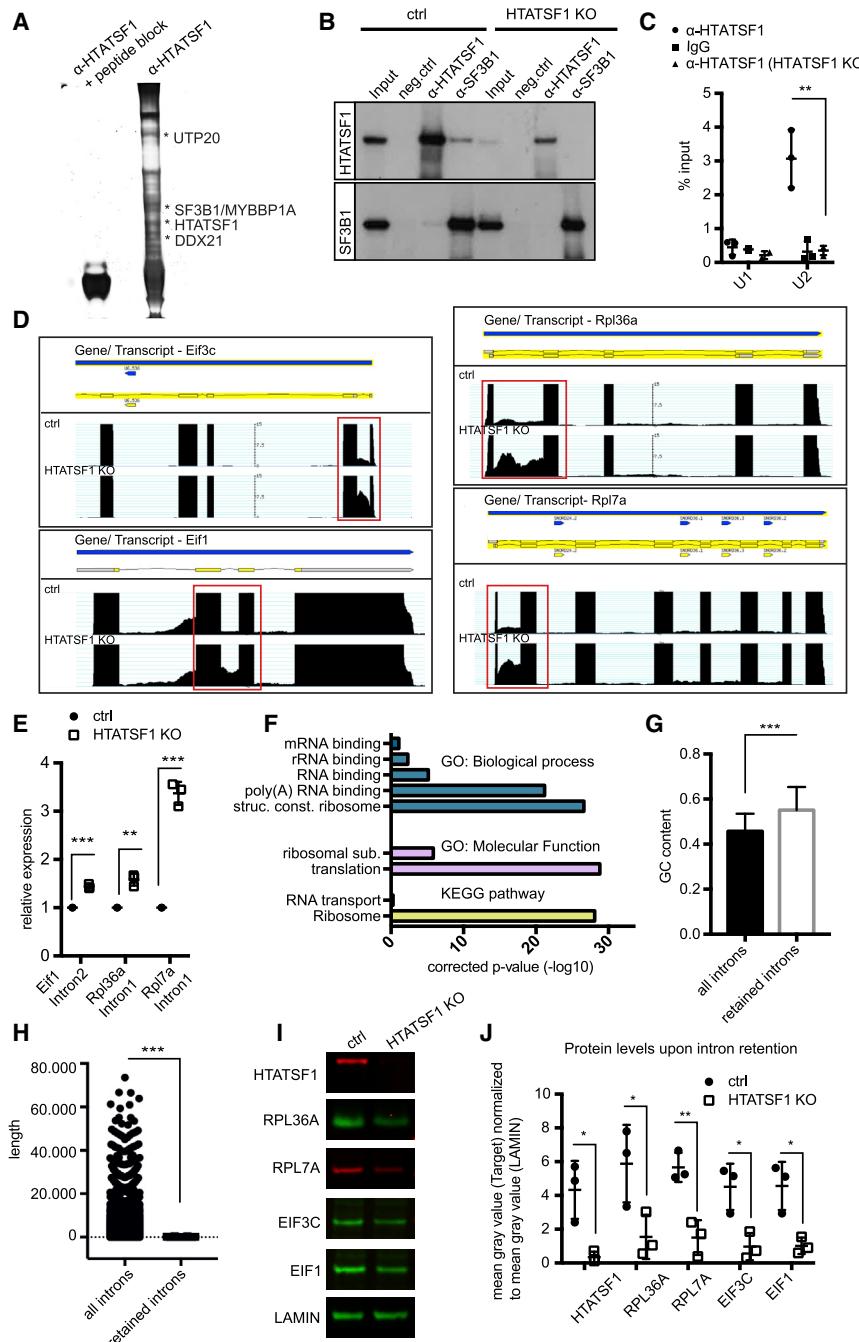


Figure 3. HTATSF1 Functions in Splicing of Ribosomal Proteins

(A) mESCs were incubated with either α -HTATSF1 antibody or α -HTATSF1 antibody and blocking peptide. Immunoprecipitates were run on a gel and silver stained. Bands were analyzed using mass spectrometry.

(B) *Htatsf1^{f/f}*, CreERT2^{+/-} ESCs were treated with 4-OH-tamoxifen or ethanol for 72 hr. Immunoprecipitation was performed with α -HTATSF1 or α -SF3B1 or non-binding ctrl antibody (neg.ctrl).

(C) *Htatsf1^{f/f}*, CreERT2^{+/-} ESCs were treated with 4-OH-tamoxifen or ethanol for 72 hr and RNA-immunoprecipitation was performed with α -HTATSF1 antibody or control immunoglobulin G (IgG) followed by qRT-PCR analysis. Values are shown as mean \pm SD; n = 3.

(D) *Htatsf1^{f/f}*, CreERT2^{+/-} ESCs were treated with 4-OH-tamoxifen or ethanol for 72 hr. RNA was subjected to RNA-seq. Examples of intron retention observed upon loss of HTATSF1 in Eif3c, Eif1, Rpl36a, and Rpl7a.

(E) *Htatsf1^{f/f}*, CreERT2^{+/-} ESCs were treated with 4-OH-tamoxifen or ethanol for 72 hr, and intron retention was quantified by qRT-PCR. Values are shown as mean \pm SD; n = 3.

(F) Gene ontology (GO) and KEGG pathway analysis of genes affected by intron retention upon loss of HTATSF1. The x axis corresponds to the negative log10 corrected p values (Benjamini).

(G) GC content of all introns and introns retained upon loss of HTATSF1.

(H) Length of all introns and introns retained upon loss of HTATSF1.

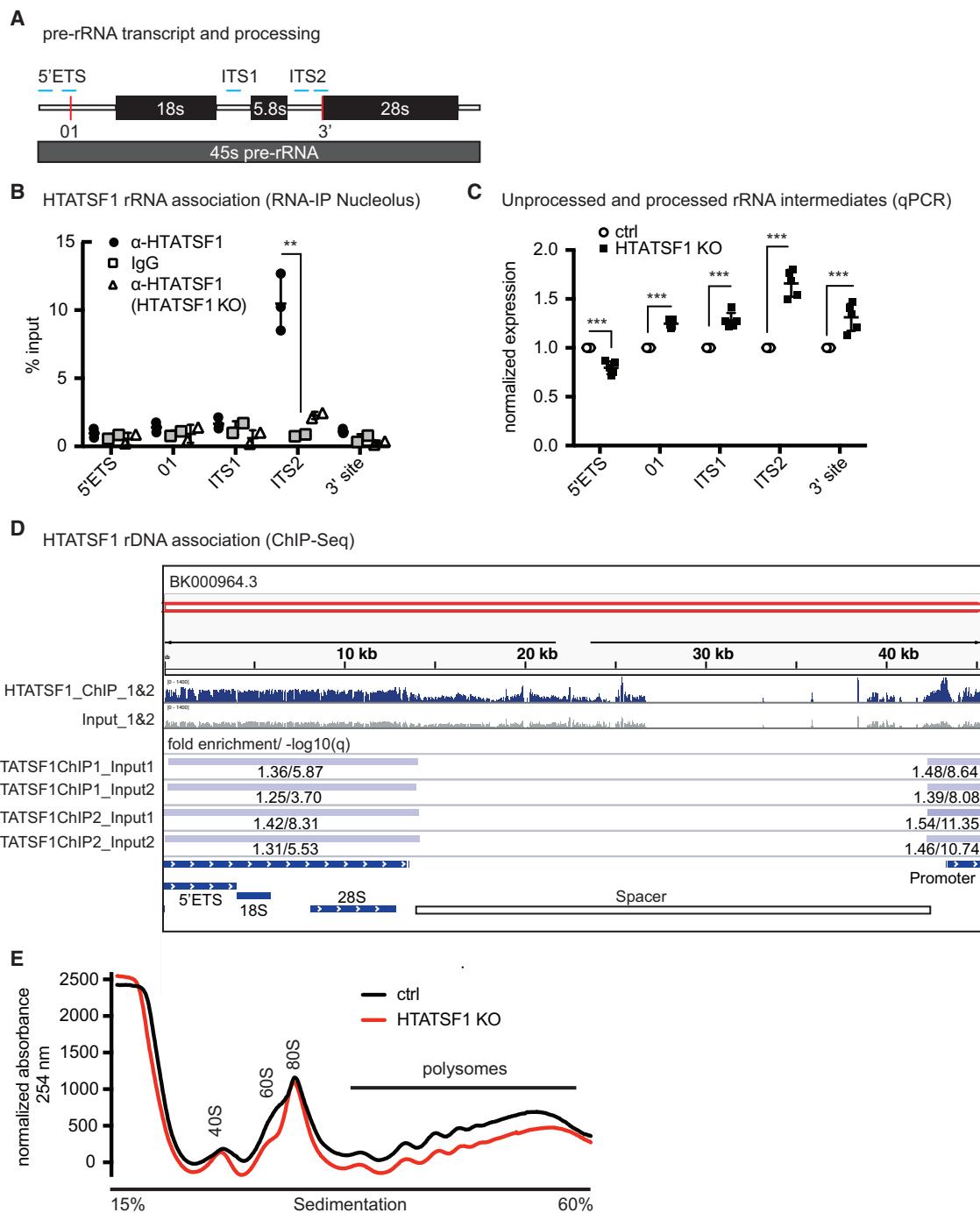
(G and H) Mann-Whitney test was used to evaluate significance. Values are shown as mean \pm SD.

(I) *Htatsf1^{f/f}*, CreERT2^{+/-} ESCs were treated with 4-OH-tamoxifen (HTATSF1 KO) or ethanol (ctrl) for 72 hr, and protein levels were analyzed by western blot. LAMIN is used as loading control.

(J) Quantification of western blot shown in (I). Mean gray value of target was normalized to mean gray value of LAMIN. Values are shown as mean \pm SD; n = 3.

and the 5.8S rRNA, which are part of the large ribosomal subunit (Figure 4A). We further analyzed levels of rRNA unprocessed transcripts and processed intermediates (Figure 4C). We observed reduction of 45S pre-rRNA levels upon HTATSF1 loss, suggesting a role in stimulation of rRNA transcription. HTATSF1 loss also resulted in an accumulation of unprocessed rRNA intermediates, most prominently in accumulation of unprocessed ITS2 containing intermediates, which underlines a role for HTATSF1 in processing of the large 60S ribosomal subunit rRNAs. Finally, ChIP-seq showed that HTATSF1 was bound predominantly to the rDNA promotor and the transcribed

region of the rDNA, although less pronounced, but not to the intergenic spacer, a binding profile similar to previously described regulators of rRNA transcription (Cong et al., 2012; Zentner et al., 2011; Calo et al., 2014) (Figure 4D). Our data indicate that HTATSF1 affects both ribosomal protein abundance through control of intron retention in ribosomal protein mRNAs as well as ribosomal RNAs through its influence on ribosomal processing and transcription. At this point, it is unclear what part of the observed phenotypes on rRNA processing and transcription is due to a decrease in ribosomal proteins production caused by intron retention in ribosomal protein mRNAs (which predominantly affects large sub-unit ribosomal proteins) and what part of the phenotype is due to a direct influence of HTATSF1 on pre-rRNA transcription and processing. It might thus be possible that a large part of the observed phenotypes is due to the splicing function of

**Figure 4. HTATSF1 Influences rRNA Transcription and Processing**

- (A) Schematic representation of unprocessed mouse rRNA transcript and selected processing sites (red). Blue indicates primers used.
- (B) *Htatsf1^{f/f}; CreERT2^{+/-}* ESCs were treated with 4-OH-tamoxifen or ethanol for 72 hr. Cells were crosslinked, and nucleoli were isolated. RNA-immunoprecipitation was performed using α-HTATSF1 antibody or IgG. Values are shown as mean ± SD; n = 3 (except for HTATSF1 KO and IgG; n = 2).
- (C) *Htatsf1^{f/f}; CreERT2^{+/-}* ESCs were treated with 4-OH-tamoxifen or ethanol for 72 hr. rRNA unprocessed transcripts and processing intermediates were quantified using qRT-PCR. Data are normalized to housekeeping gene Srp14 and to ctrl (ctrl = 1). Values are shown as mean ± SD; n = 5
- (D) ChIP was performed with α-HTATSF1 antibody and subjected to sequencing. HTATSF1 binds predominantly to the 45s pre-rRNA and promotor region of the ribosomal DNA unit. The spacer contains several repeats (SINE) which produce alignment artifacts. Fold enrichment and log10 q-value are depicted for the different replicates for ChIP (1&2) and Input (1&2) combinations for the 45s pre-rRNA and promotor region.
- (E) *Htatsf1^{f/f}; CreERT2^{+/-}* ESCs were treated with 4-OH-tamoxifen or ethanol for 72 hr. Polysomes were isolated and fractionated. Shown is 1 of 3 independent experiments showing similar results.

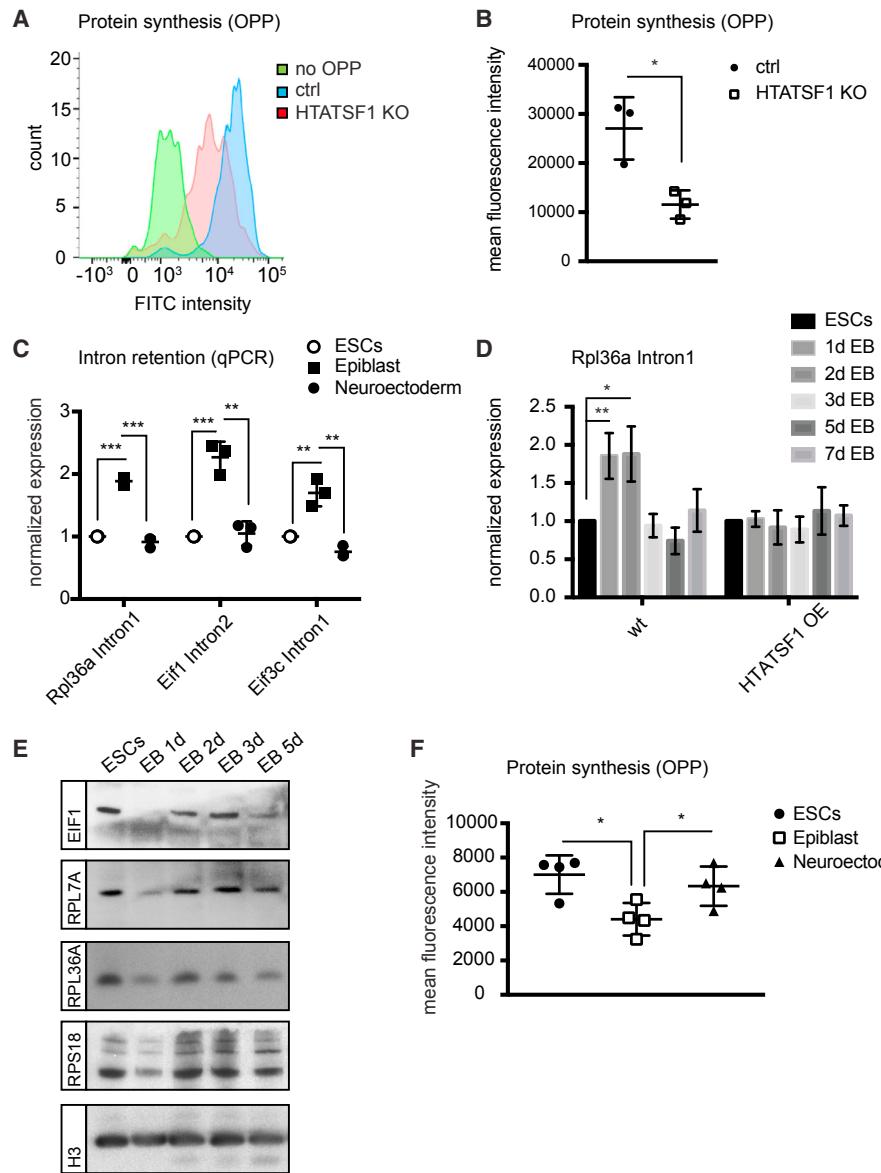


Figure 5. HTATSF1 Regulates Protein Synthesis

- (A) FACS plot of OPP incorporation in *Htatsf1*^{fl/fl}; CreERT2^{+/-} ESCs treated with 4-OH-tamoxifen or ethanol for 72 hr.
- (B) Quantification of OPP incorporation in *Htatsf1*^{fl/fl}; CreERT2^{+/-} ESCs treated with 4-OH-tamoxifen or ethanol for 72 hr. Values are shown as mean \pm SD; n = 3.
- (C) Levels of intron retention in Rpl36a, Eif1, and Eif3c in mESCs, during post-implantation epiblast differentiation, and in neuroectoderm quantified by qRT-PCR. Values are shown as mean \pm SD; n = 3.
- (D) Intron retention of Rpl36a Intron1 quantified by qRT-PCR. Values are shown as mean \pm SD; n = 3.
- (E) Data are normalized to Srp14 and to ESCs (ESCs = 1).
- (F) Western blot showing EIF1, RPL7A, RPL36A, and RPS18 protein levels during differentiation. Histone H3 was used as loading control.
- (G) OPP incorporation in ESCs, during post-implantation epiblast differentiation and in neuroectoderm. Values are shown as mean \pm SD; n = 4.

HTATSF1 rather than a direct interaction with the rRNA transcription and processing machinery.

HTATSF1 Controls Levels of the 60S Ribosomal Subunit and Protein Synthesis

We hypothesized that dual control of both ribosomal protein and rRNA levels by HTATSF1 would regulate ribosome and protein synthesis levels. We performed sucrose gradients followed by ri-

bosomal fractionation to analyze ribosome abundance (Figures 4E and S5C–S5E). Whereas levels of the 40S subunit were similar in control and HTATSF1 KO cell lines, loss of HTATSF1 led to a decrease in the 60S ribosomal subunit and a reduction in polysome levels at 72 hr after 4-OH-tamoxifen treatment. At later time points, we also observed a reduction of the 80S peak (78 hr) and of the 60S, 40S, and 80S peaks (96 hr) (Figures S5C–S5E). These results indicate that HTATSF1 predominantly

regulates 60S subunit maturation and polysome levels, but prolonged loss of HTATSF1 also affects levels of 40S and 80S. To further analyze whether HTATSF1 influences the levels of overall protein synthesis, we used O-propargyl-puromycin (OPP), a puromycin analog that generates covalent conjugates with nascent polypeptide chains. Indeed, levels of overall protein synthesis were significantly decreased upon loss of HTATSF1 (Figures 5A and 5B) confirming that HTATSF1 regulates protein synthesis levels.

HTATSF1 and Protein Synthesis Dynamics during ESC Differentiation

To test whether the HTATSF1 KO phenotype in pluripotent ESCs coincides with intron retention and reduced protein synthesis, we measured the levels of intron retention in ribosomal proteins and translation initiation factors during pluripotency and neuroectoderm differentiation by qRT-PCR (Figure 5C). We observed intron retention in HTATSF1-regulated introns of Rpl36a, Eif1, and Eif3c at the stage corresponding to the early post-implantation epiblast, where HTATSF1 expression was lowest. Overexpression of HTATSF1 abolished intron retention upon differentiation, confirming the role of HTATSF1 in intron removal (Figure 5D). In addition, EIF1, RPL7A, RPL36A, and RPS18 protein levels were high in ESCs, declined upon differentiation, and increased again in neuroectoderm (Figure 5E). Finally, OPP incorporation revealed that protein synthesis rates transiently declined in the post-implantation epiblast (Figure 5F). Taken together, our data reveal a striking correlation between HTATSF1 levels and intron retention, ribosomal protein levels, and protein translation rates in ESCs upon exit from pluripotency and differentiation into neuroectoderm.

To test the functional implications of this correlation, we attempted to rescue the HTATSF1 KO phenotype by increasing protein synthesis rates by deleting PTEN (Figure S6A) (Signer et al., 2014). *Htatsf1^{f1/f1}; CreERT2^{+/-}; Pten^{-/-}* cells had increased levels of protein synthesis (Figure S6B). We asked whether PTEN deletion in a *p53^{-/-}* background could rescue the observed loss of pluripotency upon deletion of HTATSF1. Deletion of HTATSF1 in a *p53^{-/-}, PTEN^{-/-}* background did not result in loss of NANOG, OCT4, and SOX2 (Figures 6A–6E, S6C, and S6D). In addition, loss of PTEN rescued the defect in colony-reforming efficiency observed upon loss of HTATSF1 (Figures 6F and 6G). We next assessed whether PTEN deletion could rescue failure to generate neuroectoderm upon deletion of HTATSF1 (Figure 6H). Loss of HTATSF1 prevented generation of SOX1 positive neuroectoderm in *p53^{-/-}* EBs (Figures S2L and S2M). Upon PTEN deletion, however, *PTEN^{-/-}, p53^{-/-}*, and HTATSF1 KO EBs generated SOX1 positive neuroectoderm (Figures 6I and 6J). Thus, PTEN deletion can rescue the defects observed upon loss of HTATSF1 suggesting that HTATSF1 exerts its function by controlling rates of protein synthesis.

Since loss of PTEN influences NANOG levels (Alva et al., 2011), we asked whether the rescue of the HTATSF1 KO phenotype was due to increased levels of NANOG. We overexpressed NANOG and analyzed colony-reforming efficiency upon HTATSF1 KO (Figures S6E–S6G). NANOG overexpression did not rescue the observed defect in colony-reforming efficiency, indicating that higher NANOG levels are not sufficient to rescue the HTATSF1 KO phenotype. Since PTEN also acts on AKT signaling, we asked

whether loss of HTATSF1 resulted in reduced levels of AKT Ser473 phosphorylation. We did not detect reduced levels of AKT Ser473 upon loss of HTATSF1 (Figure S6H).

In Vivo Dynamics of HTATSF1 and Protein Synthesis

To test whether the effects seen *in vitro* also occur *in vivo*, we probed early mouse embryos for expression of HTATSF1 and levels of protein synthesis. Analysis of protein synthesis via OPP incorporation and immunostaining revealed that the E4.5 epiblast had high levels of both HTATSF1 and protein synthesis (Figures 7A–7C). During development, however, levels of HTATSF1 were decreased in E5.5 and E6.5 embryos and increased again in neuroectoderm at E7.5 (Figures 7A and 7B). Protein synthesis followed HTATSF1 expression dynamics, with decreased protein synthesis observed in the post-implantation epiblast at E6.5 and an increase at E7.5 (Figures 7A and 7C). Thus, the dynamics of protein synthesis and HTATSF1 expression *in vivo* resemble the dynamics observed *in vitro* in EBs. To study the effects of loss of HTATSF1 function during early embryogenesis, we employed *in vitro* ESC and TSC stem cell (ETS) embryo technology (Harrison et al., 2017), combining GFP-labeled trophoblast stem cells (TSCs) (Tanaka et al., 1998) with *Htatsf1^{f1/f1}; CreERT2^{+/-}* ESCs. 4-OH-tamoxifen treatment deleted HTATSF1 specifically in the epiblast (Figures 7D and 7E). We first used the ETS embryo system to prevent upregulation of HTATSF1 in the post-implantation epiblast. We performed 4-OH-tamoxifen treatment 12 hr after mixing cell populations. Control ETS embryos progressed from the E4.75 stage (characterized by epiblast rosette formation) to the E5.0 stage (characterized by epiblast luminogenesis) to the E5.5 stage (characterized by trophoblast and epiblast luminogenesis) to the E5.75–E6.25 stage (characterized by proamniotic cavity fusion). Loss of HTATSF1 resulted in failure to undergo amniotic cavity fusion, indicating that HTATSF1 function is required for this process (Figure 7E). Exit from naive pluripotency is required for triggering amniotic cavity formation and developmental progression (Shahbazi et al., 2017). Accordingly, loss of the ability for cavity fusion as observed in HTATSF1 KO ETS embryos could potentially reflect failure to differentiate.

We further assessed loss of HTATSF1 during naive pluripotency in ETS embryos. We pre-treated ESCs with 4-OH-tamoxifen for 48 hr before ETS embryo generation and analyzed embryos at the time point corresponding to E4.75 (Figure S7A). While control ESCs were able to associate with TSCs to generate ETS embryos, HTATSF1 KO ESCs generated mixed aggregates with TSCs, which might indicate differentiation toward the trophoectoderm lineage, a phenotype previously described upon loss of OCT4 (Figures S7B and S7C) (Nichols et al., 1998). Taken together, these results indicate that HTATSF1 influences the regulation of morphogenesis and developmental progression during early embryonic development.

DISCUSSION

Our data identify HTATSF1 as a factor controlling early mammalian embryogenesis and pluripotency by limiting and coordinating several cell-essential processes. We propose a model where reduced rates of rRNA transcription and processing as well as mis-splicing of ribosomal mRNAs cause a reduction in

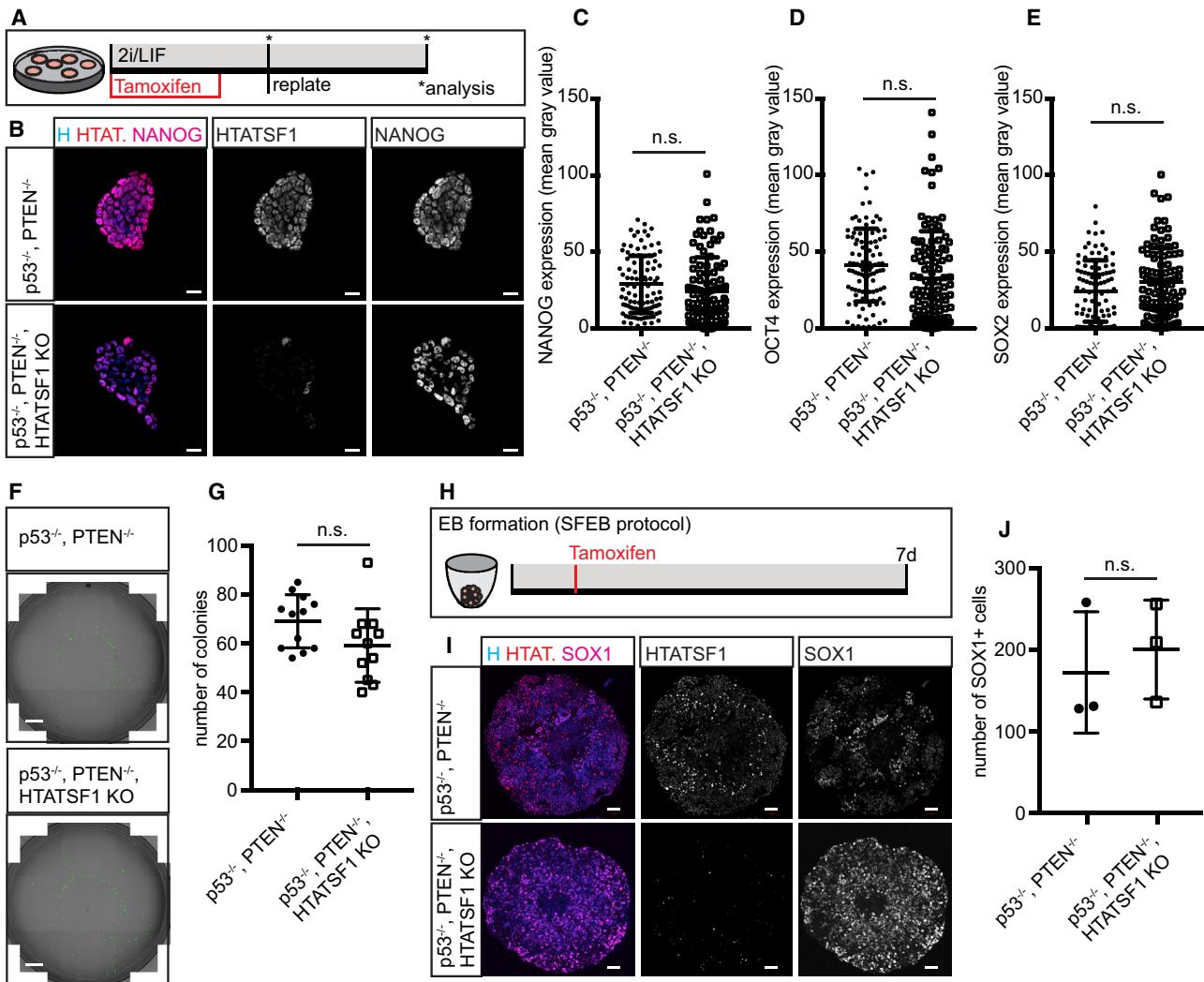


Figure 6. Loss of PTEN and p53 Can Compensate for Loss of HTATSF1

(A) Experimental workflow to assess pluripotency maintenance using immunofluorescence and colony-forming efficiency assay.

(B-E) *Htatsf1^{fl/fl}; CreERT2^{+/-}; p53^{-/-}; Pten^{-/-}* mESCs were treated with 4-OH-tamoxifen (HTATSF1 KO) or ethanol (ctrl) and analyzed 72 hr later.

(B) Immunofluorescence for HTATSF1 and NANOG. Scale bars represent 20 μ m. H, Hoechst; HTAT., HTATSF1.

(C) Quantification of NANOG expression as shown in (B).

(D) Quantification of OCT4 expression as shown in Figure S6C.

(E) Quantification of SOX2 expression as shown in Figure S6D.

(F-G) Values are shown as mean \pm SD; n = 100 cells from 3 independent experiments.

(F and G) Colony-reforming efficiency assay. *Htatsf1^{fl/fl}; CreERT2^{+/-}; p53^{-/-}; Pten^{-/-}* mESCs were treated with 4-OH-tamoxifen (HTATSF1 KO) or ethanol (ctrl) and analyzed 72 hr later. Cells were split and re-plated, and colony-reforming efficiency was analyzed.

(F) Colony formation after re-planting in p53^{-/-}, PTEN^{-/-} cells and p53^{-/-}, PTEN^{-/-}, HTATSF1 KO cells. Scale bars represent 2 mm.

(G) Quantification of colony-reforming efficiency. Values are shown as mean \pm SD; n = 10 from 4 independent experiments.

(H) Experimental workflow to assess loss of HTATSF1 function during differentiation in p53^{-/-}, PTEN^{-/-} ESCs. *Htatsf1^{fl/fl}; CreERT2^{+/-}; p53^{-/-}; Pten^{-/-}* mESCs were cultured as EBs. 24 hr after initial setup of EBs, cultures were treated with 4-OH-tamoxifen or ethanol and analyzed 7 days after treatment.

(I) Immunohistochemistry for HTATSF1 and SOX1. Scale bars represent 50 μ m. H, Hoechst; HTAT., HTATSF1.

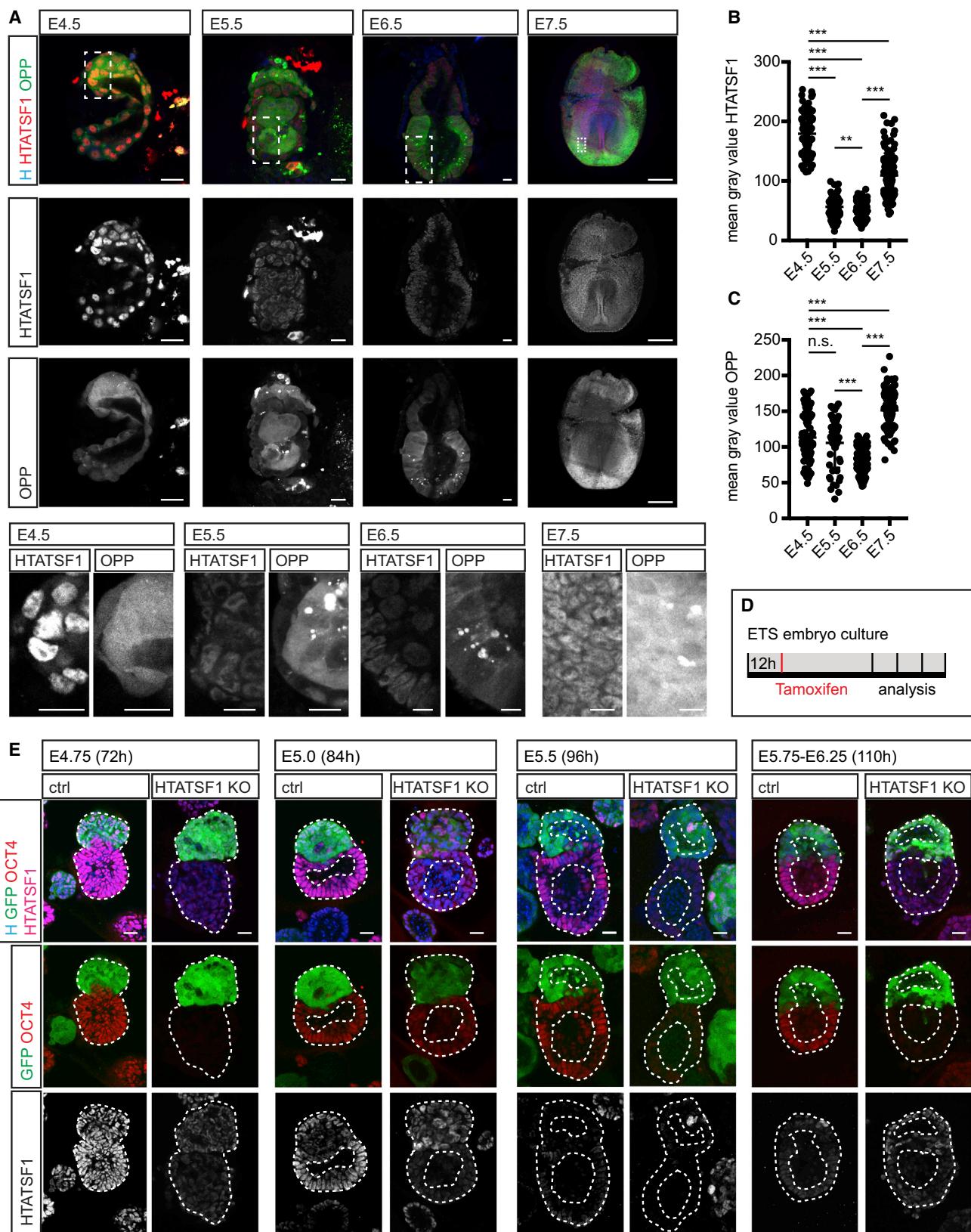
(J) Quantification of SOX1 positive cells as shown in (I). Values are shown as mean \pm SD; n = 3.

protein synthesis to facilitate the transition toward the post-implantation epiblast stage of mammalian development.

HTATSF1 Regulates Splicing

HTATSF1 was originally described as a protein that stimulates Tat-dependent elongation of HIV transcription (Zhou and Sharp, 1996) and later as a general transcriptional elongation

factor (Li and Green, 1998). It binds RNA polymerase II (Parada and Roeder, 1999) and the elongation factors P-TEFb (Fong and Zhou, 2001; Zhou et al., 1998) and DSIF/hSPT4-hSPT5 and PAF1 (Chen et al., 2009). Yeast Cus2 causes Pol II stalling and reduces transcriptional elongation when pre-spliceosome assembly is disturbed (Chatthoth et al., 2014). However, the Pol II traveling ratio in ESCs was unchanged upon HTATSF1



(legend on next page)

deletion, suggesting that the role of HTATSF1 in Pol II transcriptional elongation is not responsible for the *Htatsf1^{-/-}* ESC phenotype. Instead, our data indicate that HTATSF1 controls intron retention on a select set of genes. HTATSF1 seems specific for this type of splicing control as we could not observe changes in alternative splice site usage, exon skipping or transcript initiation. Mechanistically, we found that HTATSF1 was associated with components of the U2 snRNPs. It bound to SF3B1 that binds upstream of the branch point to allow 3' splice site recognition and is required for the splicing of most introns (Dvinge et al., 2016). Introns retained in HTATSF1 mutants had a higher than average GC content. Differences in GC content between introns and exons are crucial for spliceosomal recognition (Amit et al., 2012) and introns that have a high GC content rely on intron definition rather than exon definition for splicing (Brown et al., 2012). Thus, the specificity of HTATSF1 for ribosomal protein genes might be encoded in their intronic sequences.

HTATSF1 Regulates rRNA Transcription and Processing

We found that HTATSF1 interacts with large subunit biogenesis factors PES1 and BOP1 and influences rRNA transcription and processing. Stem cells show elevated rates of RNA transcription compared to their more differentiated daughter cells (Hayashi et al., 2014; Sanchez et al., 2016; Zhang et al., 2014). In addition, downregulation of rRNA synthesis triggers differentiation rather than being a consequence. Our results in ESCs are in line with the notion that stem cells have generally high rates of rRNA synthesis and process rRNA efficiently. We identified UTP20, a subunit of the small subunit processome as a HTATSF1 interactor. mESCs express high levels of components of the small subunit processome ensuring efficient rRNA maturation required for the maintenance of pluripotency (You et al., 2015). Another binding partner of HTATSF1 is RNA helicase DDX21, a regulator of rRNA transcription and processing (Calo et al., 2014). Thus, HTATSF1 may act in concert with rRNA regulators to both control rRNA synthesis and coordinate it with ribosomal protein synthesis.

Protein Synthesis Dynamics in Stem Cells and during Development

Our data reveal a surprising dynamic in protein synthesis rates during early mammalian development. Protein synthesis has been studied in the context of diapause, a reversible state of paused development, characterized by low levels of protein synthesis due to low MYC levels (Scognamiglio et al., 2016) and low mTOR activity (Bulut-Karslioglu et al., 2016). Besides protein synthesis, both MYC and mTOR regulate many other aspects of cellular metabolism including lipid, glucose, and nucleotide syn-

thesis (Dang, 2013; Saxton and Sabatini, 2017). Thus, paused pluripotency might be a state not only of low protein synthesis, but of overall decreased metabolism. Data in ESCs further show that loss of components of the small subunit processome, required for rRNA processing, results in reduced protein synthesis levels and loss of pluripotency maintenance in mESCs (You et al., 2015). Thus, a reduction of protein synthesis rates without affecting metabolic state leads to differentiation of mESCs. Therefore, the metabolic state of ESCs might decide whether and how levels of protein synthesis affect pluripotency maintenance. Other studies in contrast show that ESCs have low translation rates that increase during differentiation into EBs (Sampath et al., 2008; Fortier et al., 2015). Discrepancies between these findings might be due to differences in culture conditions and differentiation protocols used as well as time points analyzed.

The HTATSF1 KO phenotype could be rescued by loss of p53 and PTEN. PTEN deletion increases protein synthesis levels (Signer et al., 2014). Loss of PTEN in hESCs also results in maintenance of OCT4 and NANOG expression upon differentiation and PTEN KD hESCs and *PTEN^{-/-}* mESCs preferentially differentiate into neuroectoderm (Alva et al., 2011; Di Cristofano et al., 1998). PTEN is a major negative regulator of AKT signaling. Activation of AKT signaling in mESCs maintains pluripotency and inhibits differentiation even in the absence of LIF (Watanabe et al., 2006). Taken together, the studies on PTEN in ESCs would support a model where loss of PTEN and the accompanying increase in protein synthesis would promote pluripotency maintenance and neuroectoderm differentiation.

Reduced protein synthesis could preferentially affect proteins with low stability. Indeed, protein levels of NANOG, ESRRB, and TFCP2L1, important regulators of the naive pluripotency circuit, are rapidly reduced if translation is inhibited (You et al., 2015). Overexpression of HTATSF1 prevented differentiation and resulted in continued expression of NANOG. Thus, transiently decreasing levels of translation upon the onset of differentiation might be required to deplete proteins that confer naive pluripotency.

Control of protein synthesis by modulation of ribosome components thus emerges as a mechanism to regulate early mammalian development. Alterations in ribosome biogenesis and activity are also at the core of ribosomopathies and a hallmark of cancer cells (Truitt and Ruggero, 2016). Ribosomopathies are congenital, developmental disorders with heterogeneous clinical features ranging from developmental abnormalities to an increased cancer risk (Mills and Green, 2017; Teng et al., 2013). The rapid adjustment of levels of ribosome biogenesis through splicing of ribosomal proteins and regulation of pre-rRNA transcription and processing might thus open avenues of research not only in development but also in disease.

Figure 7. HTATSF1 and Protein Synthesis Dynamics during Mouse Embryogenesis

- (A) OPP incorporation and HTATSF1 immunohistochemistry of mouse embryos. E7.5 is a composite image. Scale bars represent 20 μm for E4.5–E6.5. Scale bar represents 200 μm for E7.5. Closeups: scale bar represents 10 μm. H, Hoechst.
- (B) Quantification of HTATSF1 expression.
- (C) Quantification of OPP incorporation.
- (B and C) Values are shown as mean ± SD; E4.5: n = 78, E5.5: n = 54, E6.5 and E7.5: n = 100 cells from 4 or 5 embryos.
- (D) Experimental workflow to assess loss of HTATSF1 in ETS embryos. *Htatsf1^{fl/fl}*; CreERT2^{+/−} mESCs and GFP-labeled TSCs were used. 4-OH-tamoxifen or ethanol was added 12 hr after ETS embryo generation.
- (E) Immunohistochemistry for OCT4 and HTATSF1. OCT4 marks ESCs, and GFP marks TSCs. Scale bars represent 20 μm. H, Hoechst.; HTAT., HTATSF1.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **CONTACT FOR REAGENT AND RESOURCE SHARING**
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHODS DETAIL**
 - Generation of the *Htatsf1* flox allele
 - Mouse and human embryoid body generation
 - EpiSC generation from mouse ESCs
 - Colony reforming efficiency assay
 - Generation of naive human ES cells from human ES cells
 - Generation of synthetic embryos (ETS embryos)
 - ChIP-qPCR and ChIP-sequencing
 - Western blots and co-IP
 - Protein synthesis measurements using OP-Puromycin Incorporation
 - Isolation of Nucleoli
 - RNA-IP
 - RNA-sequencing
 - RNA extraction and qRT-PCR
 - Plasmids
 - Polysome Extraction and analysis
 - Mass spectrometry
 - Immunohistochemistry and Immunocytochemistry
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - Analysis of ChIP-Sequencing data
 - Analysis of RNA-sequencing data
 - Analysis of mass spectrometry data
 - Statistics
- **DATA AND SOFTWARE AVAILABILITY**

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and three tables and can be found with this article online at <https://doi.org/10.1016/j.stem.2018.03.002>.

ACKNOWLEDGMENTS

N.S.C. was funded by EMBO and DFG (CO 1324/1-1) postdoctoral fellowships. Work in J.A.K.'s laboratory is supported by the Austrian Academy of Sciences, the Austrian Science Fund (grants I_1281-B19 and Z_153_B09), and an advanced grant from the European Research Council. We thank members of the Knoblich lab for helpful discussion and technical support. We thank Arabella Meixner and Michelle Foong-Sobis for the generation of *Htatsf1*^{flox} ESC lines. We thank the IMP/IMBA Biooptics Facility for FACS and help with microscopy, imaging, and image analysis. We thank the IMP/IMBA Protein Chemistry Facility, especially Karl Mechtl, Richard Imre, and Mathias Madaliński for peptide production, antibody purification, and mass spectrometry. We thank Maria Novatchkova for analysis of ChIP-seq data. We thank the Next Generation Sequencing Facility at the Vienna Biocenter Core Facilities (VBCF) for Next Generation Sequencing.

AUTHOR CONTRIBUTIONS

N.S.C. planned and performed experiments, analyzed data, and wrote the manuscript. A.M.P. and P.M. prepared samples and performed experiments. H.-C.T. isolated mouse embryos and performed blastocyst injections. M.R. performed polysome fractionation and analysis. I.M. provided resources and expertise on polysome fractionation and analysis. T.R.B. performed bioinfor-

matics analysis of RNA and ChIP-seq data. J.A.K. conceived and supervised the project, planned and interpreted experiments, and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 17, 2017

Revised: January 5, 2018

Accepted: March 7, 2018

Published: April 5, 2018

REFERENCES

- Abramczuk, M.K., Burkard, T.R., Rolland, V., Steinmann, V., Duchek, P., Jiang, Y., Wissel, S., Reichert, H., and Knoblich, J.A. (2017). The splicing co-factor Barricade/Tat-SF1 is required for cell cycle and lineage progression in *Drosophila* neural stem cells. *Development* **144**, 3932–3945.
- Aeschimann, F., Xiong, J., Arnold, A., Dieterich, C., and Grosshans, H. (2015). Transcriptome-wide measurement of ribosomal occupancy by ribosome profiling. *Methods* **85**, 75–89.
- Alva, J.A., Lee, G.E., Escobar, E.E., and Pyle, A.D. (2011). Phosphatase and tensin homolog regulates the pluripotent state and lineage fate choice in human embryonic stem cells. *Stem Cells* **29**, 1952–1962.
- Amit, M., Donyo, M., Hollander, D., Goren, A., Kim, E., Gelfman, S., Lev-Maor, G., Burstein, D., Schwartz, S., Postolsky, B., et al. (2012). Differential GC content between exons and introns establishes distinct strategies of splice-site recognition. *Cell Rep.* **1**, 543–556.
- Anders, S., Reyes, A., and Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017.
- Bao, S., Tang, F., Li, X., Hayashi, K., Gillich, A., Lao, K., and Surani, M.A. (2009). Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature* **461**, 1292–1295.
- Bell, T.J., Miyashiro, K.Y., Sul, J.-Y., Buckley, P.T., Lee, M.T., McCullough, R., Jochems, J., Kim, J., Cantor, C.R., Parsons, T.D., and Eberwine, J.H. (2010). Intron retention facilitates splice variant diversity in calcium-activated big potassium channel populations. *Proc. Natl. Acad. Sci. USA* **107**, 21152–21157.
- Blanco, S., Bandiera, R., Popis, M., Hussain, S., Lombard, P., Aleksic, J., Sajini, A., Tanna, H., Cortés-Garrido, R., Gkatza, N., et al. (2016). Stem cell function and stress response are controlled by protein synthesis. *Nature* **534**, 335–340.
- Boroviak, T., Loos, R., Bertone, P., Smith, A., and Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat. Cell Biol.* **16**, 516–528.
- Boroviak, T., Loos, R., Lombard, P., Okahara, J., Behr, R., Sasaki, E., Nichols, J., Smith, A., and Bertone, P. (2015). Lineage-specific profiling delineates the emergence and progression of naive pluripotency in mammalian embryogenesis. *Dev. Cell* **35**, 366–382.
- Boutz, P.L., Bhutkar, A., and Sharp, P.A. (2015). Detained introns are a novel, widespread class of post-transcriptionally spliced introns. *Genes Dev.* **29**, 63–80.
- Braunschweig, U., Barbosa-Morais, N.L., Pan, Q., Nachman, E.N., Alipanahi, B., Gonatopoulos-Pournatzis, T., Frey, B., Irimia, M., and Blencowe, B.J. (2014). Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res.* **24**, 1774–1786.
- Brown, S.J., Stoilov, P., and Xing, Y. (2012). Chromatin and epigenetic regulation of pre-mRNA processing. *Hum. Mol. Genet.* **21** (R1), R90–R96.
- Bryja, V., Bonilla, S., and Arenas, E. (2006). Derivation of mouse embryonic stem cells. *Nat. Protoc.* **1**, 2082–2087.
- Bulut-Karslioglu, A., Biechele, S., Jin, H., Macrae, T.A., Hejna, M., Gertsenstein, M., Song, J.S., and Ramalho-Santos, M. (2016). Inhibition of mTOR induces a paused pluripotent state. *Nature* **540**, 119–123.
- Calo, E., Flynn, R.A., Martin, L., Spitale, R.C., Chang, H.Y., and Wysocka, J. (2014). RNA helicase DDX21 coordinates transcription and ribosomal RNA processing. *Nature* **518**, 249–253.

- Chathoth, K.T., Barrass, J.D., Webb, S., and Beggs, J.D. (2014). A splicing-dependent transcriptional checkpoint associated with prespliceosome formation. *Mol. Cell* 53, 779–790.
- Chen, Y., Yamaguchi, Y., Tsugeno, Y., Yamamoto, J., Yamada, T., Nakamura, M., Hisatake, K., and Handa, H. (2009). DSIF, the Paf1 complex, and Tat-SF1 have nonredundant, cooperative roles in RNA polymerase II elongation. *Genes Dev.* 23, 2765–2777.
- Cong, R., Das, S., Ugrinova, I., Kumar, S., Mongelard, F., Wong, J., and Bouvet, P. (2012). Interaction of nucleolin with ribosomal RNA genes and its role in RNA polymerase I transcription. *Nucleic Acids Res.* 40, 9441–9454.
- Dang, C.V. (2013). MYC, metabolism, cell growth, and tumorigenesis. *Cold Spring Harb. Perspect. Med.* Published online August 1, 2013. <https://doi.org/10.1101/cshperspect.a014217>.
- Di Cristofano, A., Pesce, B., Cordon-Cardo, C., and Pandolfi, P.P. (1998). Pten is essential for embryonic development and tumour suppression. *Nat. Genet.* 19, 348–355.
- Dvinge, H., Kim, E., Abdel-Wahab, O., and Bradley, R.K. (2016). RNA splicing factors as oncoproteins and tumour suppressors. *Nat. Rev. Cancer* 16, 413–430.
- Fong, Y.W., and Zhou, Q. (2001). Stimulatory effect of splicing factors on transcriptional elongation. *Nature* 414, 929–933.
- Fortier, S., MacRae, T., Bilodeau, M., Sargeant, T., and Sauvageau, G. (2015). Haploinsufficiency screen highlights two distinct groups of ribosomal protein genes essential for embryonic stem cell fate. *Proc. Natl. Acad. Sci. USA* 112, 2127–2132.
- Harrison, S.E., Sozen, B., Christodoulou, N., Kyriyanou, C., and Zernicka-Goetz, M. (2017). Assembly of embryonic and extraembryonic stem cells to mimic embryogenesis in vitro. *Science*. Published online April 14, 2017. <https://doi.org/10.1126/science.aal1810>.
- Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* 146, 519–532.
- Hayashi, Y., Kuroda, T., Kishimoto, H., Wang, C., Iwama, A., and Kimura, K. (2014). Downregulation of rRNA transcription triggers cell differentiation. *PLoS ONE* 9, e98586.
- Hochstatter, J., Hözel, M., Rohrmoser, M., Schermelleh, L., Leonhardt, H., Keough, R., Gonda, T.J., Imhof, A., Eick, D., Längst, G., and Németh, A. (2012). Myb-binding protein 1a (Mybbp1a) regulates levels and processing of pre-ribosomal RNA. *J. Biol. Chem.* 287, 24365–24377.
- Huang, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13.
- Huang, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.
- Ivics, Z., Hackett, P.B., Plasterk, R.H., and Izsvák, Z. (1997). Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* 91, 501–510.
- Kalkan, T., Olova, N., Roode, M., Mulas, C., Lee, H.J., Nett, I., Marks, H., Walker, R., Stunnenberg, H.G., Lilley, K.S., et al. (2017). Tracking the embryonic stem cell transition from ground state pluripotency. *Development* 144, 1221–1234.
- Kamiya, D., Banno, S., Sasai, N., Ohgushi, M., Inomata, H., Watanabe, K., Kawada, M., Yakura, R., Kiyonari, H., Nakao, K., et al. (2011). Intrinsic transition of embryonic stem-cell differentiation into neural progenitors. *Nature* 470, 503–509.
- Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S., and Karolchik, D. (2010). BigWig and BigBed: Enabling browsing of large distributed datasets. *Bioinformatics* 26, 2204–2207.
- Kojima, Y., Kaufman-Francis, K., Studdert, J.B., Steiner, K.A., Power, M.D., Loebel, D.A.F., Jones, V., Hor, A., de Alencastro, G., Logan, G.J., et al. (2014). The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. *Cell Stem Cell* 14, 107–120.
- Lancaster, M.A., Renner, M., Martin, C.-A., Wenzel, D., Bicknell, L.S., Hurles, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. *Nature* 501, 373–379.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760.
- Li, X.-Y., and Green, M.R. (1998). The HIV-1 Tat cellular coactivator Tat-SF1 is a general transcription elongation factor. *Genes Dev.* 12, 2992–2996.
- Li, Z.F., and Lam, Y.W. (2014). A new rapid method for isolating nucleoli. In *Methods in Molecular Biology*, R. Hancock, ed. (Springer), pp. 35–42.
- Lin, T., Chao, C., Saito, S., Mazur, S.J., Murphy, M.E., Appella, E., and Xu, Y. (2004). p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. *Nat. Cell Biol.* 7, 165–171.
- Mátés, L., Chuah, M.K.L., Belay, E., Jerchow, B., Manoj, N., Acosta-Sánchez, A., Grzela, D.P., Schmitt, A., Becker, K., Matrai, J., et al. (2009). Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nat. Genet.* 41, 753–761.
- Matsuda, T., and Cepko, C.L. (2004). Electroporation and RNA interference in the rodent retina in vivo and in vitro. *Proc. Natl. Acad. Sci. USA* 101, 16–22.
- Mauger, O., Lemoine, F., and Scheiffele, P. (2016). Targeted intron retention and excision for rapid gene regulation in response to neuronal activity. *Neuron* 92, 1266–1278.
- Mills, E.W., and Green, R. (2017). Ribosomopathies: There's strength in numbers. *Science* 358. Published online November 3, 2017. <https://doi.org/10.1126/science.aan2755>.
- Nakashima, T., Hayashi, M., Fukunaga, T., Kurata, K., Oh-hora, M., Feng, J.Q., Bonewald, L.F., Kodama, T., Wutz, A., Wagner, E.F., et al. (2011). Evidence for osteocyte regulation of bone homeostasis through RANKL expression. *Nature Medicine* 17, 1231–1234.
- Neumüller, R.A., Richter, C., Fischer, A., Novatchkova, M., Neumüller, K.G., and Knoblich, J.A. (2011). Genome-wide analysis of self-renewal in Drosophila neural stem cells by transgenic RNAi. *Cell Stem Cell* 8, 580–593.
- Nichols, J., and Smith, A. (2012). Pluripotency in the embryo and in culture. *Cold Spring Harb. Perspect. Biol.* 4, a008128–a008128.
- Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 95, 379–391.
- Noormohammadi, A., Calculli, G., Gutierrez-Garcia, R., Khodakarami, A., Koyuncu, S., and Vilchez, D. (2018). Mechanisms of protein homeostasis (proteostasis) maintain stem cell identity in mammalian pluripotent stem cells. *Cell. Mol. Life Sci.* 75, 275–290.
- Parada, C.A., and Roeder, R.G. (1999). A novel RNA polymerase II-containing complex potentiates Tat-enhanced HIV-1 transcription. *EMBO J.* 18, 3688–3701.
- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842.
- Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McGuire, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. *Cell* 141, 432–445.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308.
- Sampath, P., Pritchard, D.K., Pabon, L., Reinecke, H., Schwartz, S.M., Morris, D.R., and Murry, C.E. (2008). A hierarchical network controls protein translation during murine embryonic stem cell self-renewal and differentiation. *Cell Stem Cell* 2, 448–460.
- Sanchez, C.G., Teixeira, F.K., Czech, B., Preall, J.B., Zamparini, A.L., Seifert, J.R.K., Malone, C.D., Hannon, G.J., and Lehmann, R. (2016). Regulation of ribosome biogenesis and protein synthesis controls germline stem cell differentiation. *Cell Stem Cell* 18, 276–290.

- Saxton, R.A., and Sabatini, D.M. (2017). mTOR signaling in growth, metabolism, and disease. *Cell* 168, 960–976.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: An open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682.
- Scognamiglio, R., Cabezas-Wallscheid, N., Thier, M.C., Altamura, S., Reyes, A., Prendergast, Á.M., Baumgärtner, D., Carnevalli, L.S., Atzberger, A., Haas, S., et al. (2016). Myc depletion induces a pluripotent dormant state mimicking diapause. *Cell* 164, 668–680.
- Shahbazi, M.N., Scialdone, A., Skorupska, N., Weberling, A., Recher, G., Zhu, M., Jedrusik, A., Devito, L.G., Noli, L., Macaulay, I.C., et al. (2017). Pluripotent state transitions coordinate morphogenesis in mouse and human embryos. *Nature* 552, 239–243.
- Signer, R.A.J., Magee, J.A., Salic, A., and Morrison, S.J. (2014). Haematopoietic stem cells require a highly regulated protein synthesis rate. *Nature* 509, 49–54.
- Smith, A. (2017). Formative pluripotency: The executive phase in a developmental continuum. *Development* 144, 365–373.
- Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., et al. (2014). Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* 158, 1254–1269.
- Tanaka, S., Kunath, T., Hadjantonakis, A.K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. *Science* 282, 2072–2075.
- Taus, T., Köcher, T., Pichler, P., Paschke, C., Schmidt, A., Henrich, C., and Mechtlar, K. (2011). Universal and confident phosphorylation site localization using phosphoRS. *J. Proteome Res.* 10, 5354–5362.
- Teng, T., Thomas, G., and Mercer, C.A. (2013). Growth control and ribosomopathies. *Curr. Opin. Genet. Dev.* 23, 63–71.
- Tosolini, M., and Jouneau, A. (2015). From naive to primed pluripotency: In vitro conversion of mouse embryonic stem cells in epiblast stem cells. In *Methods in Molecular Biology* (Springer), pp. 209–216.
- Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: Discovering splice junctions with RNA-seq. *Bioinformatics* 25, 1105–1111.
- Truitt, M.L., and Ruggero, D. (2016). New frontiers in translational control of the cancer genome. *Nat. Rev. Cancer* 16, 288–304.
- Wang, L., Wang, S., and Li, W. (2012). RSeQC: Quality control of RNA-seq experiments. *Bioinformatics* 28, 2184–2185.
- Watanabe, S., Umehara, H., Murayama, K., Okabe, M., Kimura, T., and Nakano, T. (2006). Activation of Akt signaling is sufficient to maintain pluripotency in mouse and primate embryonic stem cells. *Oncogene* 25, 2697–2707.
- Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J.H. (2016). Dynamic stem cell states: Naive to primed pluripotency in rodents and humans. *Nat. Rev. Mol. Cell Biol.* 17, 155–169.
- Wong, J.J.L., Ritchie, W., Ebner, O.A., Selbach, M., Wong, J.W.H., Huang, Y., Gao, D., Pinello, N., Gonzalez, M., Baidya, K., et al. (2013). Orchestrated intron retention regulates normal granulocyte differentiation. *Cell* 154, 583–595.
- You, K.T., Park, J., and Kim, V.N. (2015). Role of the small subunit processome in the maintenance of pluripotent stem cells. *Genes Dev.* 29, 2004–2009.
- Zentner, G.E., Saiakova, A., Manaenkov, P., Adams, M.D., and Scacheri, P.C. (2011). Integrative genomic analysis of human ribosomal DNA. *Nucleic Acids Res.* 39, 4949–4960.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-seq (MACS). *Genome Biol.* 9, R137.
- Zhang, Q., Shalaby, N.A., and Buszczak, M. (2014). Changes in rRNA transcription influence proliferation and cell fate within a stem cell lineage. *Science* 343, 298–301.
- Zhou, Q., and Sharp, P.A. (1996). Tat-SF1: Cofactor for stimulation of transcriptional elongation by HIV-1 Tat. *Science* 274, 605–610.
- Zhou, Q., Chen, D., Pierstorff, E., and Luo, K. (1998). Transcription elongation factor P-TEFb mediates Tat activation of HIV-1 transcription at multiple stages. *EMBO J.* 17, 3681–3691.

STAR★METHODS**KEY RESOURCES TABLE**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Sox1	R&D Systems	Cat#AF3369; RRID:AB_2239879
N-Cadherin	BD	Cat#610920; RRID:AB_2077527
E-Cadherin	BD	Cat#610182; RRID:AB_397581
Sf3b1	Acris Antibodies	Cat#AM26528AF-N; RRID:AB_2721250
DDX 21(F-5)	Santa Cruz	Cat#sc-376758; RRID:AB_2721251
Ribosomal Protein L7A (H-266)	Santa Cruz	Cat#sc-98618; RRID:AB_2238683
Ribosomal Protein S18 (G-13)	Santa Cruz	Cat#sc-161199; RRID:AB_2180186
Ribosomal Protein L36a (43-A)	Santa Cruz	Cat#sc-100831; RRID:AB_2238666
Oct-3/4 (C-10)	Santa Cruz	Cat#sc-5279; RRID:AB_628051
MYBBP1A (F-25)	Santa Cruz	Cat#sc-133800; RRID:AB_10611660
Lamin A/C (4C11)	Cell Signaling	Cat#4777S; RRID:AB_10545756
Histone H3 (D2B12) XP-R	Cell Signaling	Cat#4620S; RRID:AB_1904005
eIF1 (D7G3L)	Cell Signaling	Cat#12496S; RRID:AB_2721252
p53 (PAb 240)	Abcam	Cat#ab26; RRID:AB_303198
Nanog	Abcam	Cat#ab80892; RRID:AB_2150114
Nanog-ChIP Grade	Abcam	Cat#ab21624; RRID:AB_446437
Nanog	R&D Systems	Cat#AF2729; RRID:AB_2150103
htatsf1 (rb) (raised in rb against mouse SGNNLSGNDEFDEQLR-C)	This paper	N/A
htatsf1 (gp) (raised in gp against mouse SGNNLSGNDEFDEQLR-C)	This paper	N/A
BOP 1	Abcam	Cat#ab86652; RRID:AB_1924815
Pescadillo (H-10)	Santa Cruz	Cat#sc-166300; RRID:AB_2252171
Akt	Cell Signaling	Cat#9272S; RRID:AB_329827
P-Akt (Ser473)	Cell Signaling	Cat#4060S; RRID:AB_2315049
Pol II N-20	Santa Cruz	Cat#sc-899; RRID:AB_632359
Brachyury	R&D Systems	Cat#AF2085; RRID:AB_2303014
Sox2	R&D Systems	Cat#MAB2018; RRID:AB_358009
Sox17	R&D Systems	Cat#AF1924; RRID:AB_355060
Eif3c	Cell Signaling	Cat#2068S; RRID:AB_2096742
Fibrillarin (rabbit)	Abcam	Cat#ab5821; RRID:AB_2105785
Fibrillarin (mouse)	Abcam	Cat#ab4566; RRID:AB_304523
PTEN (13866)	Cell Signaling	Cat#9559S; RRID:AB_390810
Chemicals, Peptides, and Recombinant Proteins		
Hoechst 33258 (bis-benzimide)	ThermoFisher Scientific	Cat#H-3569
Trizol	ThermoFisher Scientific	Cat#15596018
SuperScript III Reverse Transcriptase	Invitrogen	Cat#18080-044
SGNNLSGNDEFDEQLR-C peptide sequence	This paper	N/A
htatsf1 antibody		
Gelatin	Sigma-Aldrich	Cat#G1890
Neurobasal medium	GIBCO	Cat#21103-049
DMEM	GIBCO	Cat#11960-044
Penicillin/Streptomycin	GIBCO	Cat#15140-122
DMEM/F12	GIBCO	Cat#11320-074
Knockout Serum Replacement	GIBCO	Cat #10828-028

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
N2-supplement	Thermo Fisher Scientific	Cat#17502001
B27-supplement	Thermo Fisher Scientific	Cat#17504044
FBS	GIBCO	Cat#10270-106
BSA	VWR International	Cat#422351S
Glutamax	Thermo Fisher Scientific	Cat#35050-038
MEM-NEAA	Sigma-Aldrich	Cat#M7145
β-Mercaptoethanol	Sigma-Aldrich	Cat#60-24-2
ESGRO mouse LIF	Millipore	Cat#ESG1107
CHIR99021	Calbiochem	Cat#361571
PD0325901	Calbiochem	Cat#444968
Sodium Pyruvate	Sigma-Aldrich	Cat#P2256
4 OH-Tamoxifen	Sigma-Aldrich	Cat#H7904
GMEM	Sigma-Aldrich	Cat#G5154
KOSR	GIBCO	Cat#10828028
mTeSR1	Stem cell technologies	Cat#05851
Fibronectin	Sigma-Aldrich	Cat#F0895
ActivinA	R&D Systems	Cat#338-AC-050
Collagenase	GIBCO	Cat#17104019
Matrigel	Corning	Cat#354234
Accutase	Sigma-Aldrich	Cat#A6964
Y-27632 (Rock Inhibitor)	Calbiochem	Cat#688000
Human LIF	Peprotech	Cat#AF-300-05
Gö6983	Sigma-Aldrich	Cat#G1918
Na-deoxycholate	ThermoFisher Scientific	Cat#89904
NP-40 substitute	Sigma-Aldrich	Cat#74385
GlycoBlue	Ambion	Cat#AM9515
Complete protease inhibitor cocktail	Roche	Cat#05 056489001
PhosphoStop	Roche	Cat#04 906837001
O-propargyl-puromycin	Jenabioscience	Cat#NV-931-05
M2 medium	Sigma-Aldrich	Cat#M7167
EmbryoMax KSOM medium	Millipore	Cat#MR-020P-5F
RNase inhibitor SUPERaseIN	Ambion	Cat#AM2694
SuperScript III Reverse Transcriptase	ThermoFisher Scientific	Cat#18080044
Actinomycin D	Sigma-Aldrich	Cat#A1410
Cycloheximide	Sigma-Aldrich	Cat#C1988
2-Methylbutane	Sigma-Aldrich	Cat#32631
Fluorescent Mounting Medium	Dako	Cat#S3023
bFGF	In house	N/A
Trypsin-EDTA	GIBCO	Cat#25300054
Heparin sodium salt	Sigma-Aldrich	Cat#H3149-10KU
RPMI 1640 medium	Sigma-Aldrich	Cat#M3817
ETS-embryo medium	Cell guidance systems	Cat#M13
FGF-4	Peprotech	Cat#100-31
L-glutamine	GIBCO	Cat#25030
G418	Invivogen	Cat#ant-gn-1
RNasin	Promega	Cat#N2515
Critical Commercial Assays		
SimpleCHIP Enzymatic Chromatin IP Kit	Cell Signaling Technology	Cat#9003
Human Stem Cell Nucleofector Kit 1	Amaxa	Cat#VVPH-5012

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Amaxa Mouse ES Cell Nucleofector Kit	Lonza	Cat#VPH-1001
Click-iT Plus OPP Alexa Fluor 488 Protein Synthesis Assay Kit	ThermoFisher Scientific	Cat#C10456
Click-iT Plus OPP Alexa Fluor 647 Protein Synthesis Assay Kit	ThermoFisher Scientific	Cat#C10458
Vector Blue AP substrate kit	Vector Laboratories	Cat#SK-5300
Kapa HTP Library Preparation Kit	Roche	Cat#07961901001
Western blot development kit	GE Healthcare	Cat#RPN 2106
Dynabeads mRNA purification kit	ThermoFisher Scientific	Cat#61006
Turbo DNA-Free kit	ThermoFisher Scientific	Cat#AM1907
Pierce BCA Protein Assay Kit	ThermoFisher Scientific	Cat#23227
Pierce Silver Stain Kit	ThermoFisher Scientific	Cat#24600
Infinium_PsychArray_24_v1.1	Illumina	Cat#20015239
Deposited Data		
Chip-Seq data PolII	GEO: GSE96595	https://www.ncbi.nlm.nih.gov/geo/
Chip-Seq data HTATSF1	GEO: GSE96595	https://www.ncbi.nlm.nih.gov/geo/
RNA sequencing data	GEO: GSE96595	https://www.ncbi.nlm.nih.gov/geo/
Experimental Models: Cell Lines		
mESCs (HTATSF1 ^{f/f} CreERT2 ^{+/−})	This paper	N/A
mouse embryonic fibroblasts (MEF CF-1 IRR)	MTI-GlobalStem	Cat#GSC-6001G; CVCL_K232
H9 human embryonic stem cells	WiCell	Cat#WA09-DL-11 p26; RRID: CVCL_9773
N2A mouse neuroblastoma cell line	ATCC	ATCC CCL-131; RRID: CVCL_0470
A9 mouse ESC line (C57/BL6J × 129/Sv ES)	Nakashima et al., 2011	N/A
TS-GFP	Tanaka et al., 1998	N/A
Experimental Models: Organisms/Strains		
Mouse:C57/BL6J	In house bred	N/A
Mouse: HTATSF1 ^{f/f}	This paper	N/A
Mouse: Gt(ROSA)26Sor ^{tm1(cre/ERT2)Ty/J}	Jackson Laboratories	Cat#008463
Oligonucleotides		
Genomic PCR Primers, see Table S3	N/A	N/A
Primers used for qPCR, see Table S3	N/A	N/A
Recombinant DNA		
pCMV(CAT)T7-SB100	Addgene	Cat#34879
pCAGEN	Addgene	Cat#11160
pT2/LTR7-GFP	Addgene	Cat#62541
pCAG-GS/IR	This paper	N/A
pSpCas9n(BB)-2A-Puro (PX462) V2.0	Addgene	Cat#62987
pSpCas9n(BB)-2A-GFP (PX461)	Addgene	Cat#48140
pCAG-hNanog-p2A-eGFP	This paper	N/A
pCAG-hKLF2-p2A-eGFP	This paper	N/A
pCAG-HTATSF1-p2A-eGFP	This paper	N/A
Software and Algorithms		
Celigo Software	Nexcelom	N/A
Bowtie	Langmead et al., 2009	http://bowtie-bio.sourceforge.net/index.shtml
bigWigs	Kent et al., 2010	N/A
BigBed	Kent et al., 2010	N/A
Fiji	Schindelin et al., 2012	https://fiji.sc/
BWA (v0.6.1)	Li and Durbin, 2009	http://bio-bwa.sourceforge.net/

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
TopHat (v1.4.1)	Trapnell et al., 2009	http://ccb.jhu.edu/software/tophat/index.shtml
DEXSeq	Bioconductor	http://bioconductor.org/packages/release/bioc/html/DEXSeq.html
DAVID Bioinformatics Analyses Tool	Huang et al., 2009b	https://david.ncifcrf.gov/
Proteome Discoverer (v1.3.0.339)	ThermoFisher Scientific	https://www.thermofisher.com/order/catalog/product/OPTON-30795
Mascot (v2.2.07)	Matrix Science, London, UK	http://www.matrixscience.com/
phosphoRS	Taus et al., 2011	http://ms.imp.ac.at/?goto=phosphors
Scaffold (v3.3.1)	Proteome Software	http://www.proteomesoftware.com/products/scaffold/
GraphPad Prism version 7	GraphPad Software, La Jolla California, USA	https://www.graphpad.com/scientific-software/prism/
picard-tools (1.82)	N.A.	http://broadinstitute.github.io/picard
bedtools	Quinlan and Hall, 2010	http://bedtools.readthedocs.io/en/latest/
Macs 1.4	Zhang et al., 2008	https://pypi.python.org/pypi/MACS/1.4.3
Other		
Ultra-low-cluster 96-well plate	Szabo-Scandic	Cat#COR7007
RNA 6000 Nano Chip	Agilent Technologies	Cat#5067-1511
Costar Ultra-low cluster, 24-well plates	Szabo-Scandic	Cat#COS3473
NuPage 4%–12% Bis-Tris Gel	Invitrogen	Cat#NP0321BOX
Odyssey Nitrocellulose Membrane	LI-COR Biosciences	Cat#926-31092
Dynabeads protein A	Thermo Fisher Scientific	Cat#10002D
rmp protein A Sepharose fast flow	GE Healthcare	Cat#17-5138-01

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Juergen A. Knoblich (juergen.knoblich@imba.oewa.ac.at).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

6- to 12-week old mice were used throughout the study. C57/BL6J mice and *Htatsf1*^{f/f} mice were bred in house in the animal facility of IMBA, Vienna, Austria. All mice were maintained according to the ethical animal license protocol, complying with Austrian and EU legislation. Mice were housed in a controlled environment with a 12:12 hour light-dark cycle with food and water provided *ad libitum* under specific pathogen free (SPF) conditions. Mice had access to shelters and nesting material. All experiments were approved by the Federal Ministry for Science and Research, Vienna, Austria.

All mouse cell lines were cultured at 37°C in a controlled humidified atmosphere of 5% CO₂. All mouse ESCs used routinely tested negative for mycoplasma. Mouse ESCs were grown on Gelatin (Sigma-Aldrich cat. #G1890)-coated 25 cm² bottles in 2i/LIF medium containing Neurobasal medium (GIBCO cat. #21103-049), DMEM/F12 (GIBCO cat. #11320-074), N2-supplement (Thermo Fisher Scientific cat. #17502001), B27-supplement (Thermo Fisher Scientific cat. #17504044), 2% FBS (tested for stem cell culture) (GIBCO cat. #10270-106), 0.05% BSA (VWR International cat. #422351S), Glutamax (Thermo Fisher Scientific cat. #35050-038), 1% MEM-NEAA (Sigma-Aldrich cat. #M7145), 0.1 mM β-Mercaptoethanol (Sigma-Aldrich cat. #60-24-2). Additionally, the medium was supplemented freshly with 10³ U/mL LIF (Millipore cat. #ESG1107), 3 μM CHIR99021 (Calbiochem cat. #361571) and 1 μM PD0325901 (Calbiochem cat. #444968). Cells were split using 0.025% Trypsin-EDTA (GIBCO cat. #25300054). For 2i/LIF withdrawal experiments, the aforementioned medium was used without addition of CHIR99021 (Calbiochem cat. #361571), PD0325901 (Calbiochem cat. #444968) and LIF (Millipore cat. #ESG1107). Mouse ES cells were nucleofected using Amaxa Mouse ES Cell Nucleofector Kit (Lonza cat. #VPH-1001) according to the manufacturer's instructions. Cells were detached using 0.025% Trypsin-EDTA (GIBCO cat. #25300054). We used 1x 10⁶ cells per nucleofection. After nucleofection, cells were seeded into gelatin-coated 24-well plates and further expanded. Constructs for overexpression were based on the Sleeping Beauty Transposase system. Constructs included p2A_eGFP for identification of positive cells. Positive cells were purified using Fluorescence Associated Cell Sorting (FACS). Deletion of p53 and PTEN using CRISPR/Cas9 was performed as described (Ran et al., 2013).

Mouse ESCs were grown on MEFs in conventional ESC medium containing DMEM (GIBCO cat. #11960-044), 15% FBS (GIBCO cat. #10270-106), Glutamax (Thermo Fisher Scientific cat. #35050-038), 1% Sodium Pyruvate, 1% MEM-NEAA (Sigma-Aldrich cat. #M7145), 0.1 mM β-Mercaptoethanol (Sigma-Aldrich cat. #60-24-2) and 10³ U/mL ESGRO mouse LIF (Millipore cat. #ESG1107). Cells were split using 0.05% Trypsin-EDTA (GIBCO cat. #25300054). Gender of mouse hybrid C57/Bl6 x 129/SV A9 ESCs has not been determined.

Mouse *HTATSF1*^{f/f}, CreERT2^{+/−} and *HTATSF1*^{f/f} as well as *HTATSF1*^{f/+}; CreERT2^{+/−} ESCs were derived from E3.5 blastocysts as previously described (Bryja et al., 2006). Briefly, female mice were paired with males in the evening and on the following morning examined for vaginal plugs. Females with positive vaginal plugs were considered as 0.5 days post-coitum (dpc) or embryonic day 0.5 (E0.5). E3.5 blastocysts were flushed from the oviducts with DMEM (GIBCO cat. #11960-044) medium and plated onto MEF feeder cells in medium containing DMEM (GIBCO cat. #11960-044) supplemented with 20% Knockout Serum Replacement (GIBCO cat. #10828-028), penicillin (100 U ml^{−1})/streptomycin (100 µg ml^{−1}) (GIBCO cat. #15140-122), 2 mM L-glutamine (GIBCO cat. #25030), 1 × MEM-NEAA (Sigma-Aldrich cat. #M7145), 100 µM β-Mercaptoethanol (Sigma-Aldrich cat. #60-24-2) and 10³ U/mL of recombinant ESGRO mouse LIF (Millipore cat. #ESG1107). After attachment and hatching of blastocysts, ICM outgrowths were passaged using Trypsin-EDTA (GIBCO cat. #25300054) and grown further in DMEM (GIBCO cat. #11960-044) containing 15% FBS (GIBCO cat. #10270-106), Glutamax (Thermo Fisher Scientific cat. #35050-038), 1% Sodium Pyruvate (Sigma-Aldrich cat. #P2256), 1% MEM-NEAA (Sigma-Aldrich cat. #M7145), 0.1 mM β-Mercaptoethanol (Sigma-Aldrich cat. #60-24-2) and 10³ U/mL ESGRO mouse LIF (Millipore cat. #ESG1107). Deletion of *HTATSF1* was performed by treatment of *HTATSF1*^{f/f}; CreERT2^{+/−} ESCs with 4-OH-Tamoxifen (800 nM) (Sigma-Aldrich cat. #H7904). 4-OH-Tamoxifen was solubilized in Ethanol. Gender of mouse *HTATSF1*^{f/f}, CreERT2^{+/−} and *HTATSF1*^{f/f} as well as *HTATSF1*^{f/+}; CreERT2^{+/−} ESCs has not been determined.

All human cell lines were cultured at 37°C in a controlled humidified atmosphere of 5% CO₂. Human ES cell lines used were H9 (WiCell), which have been authenticated by the provider and routinely tested negative for mycoplasma. Genome integrity and gender (female) were confirmed using Infinium_PsychArray_24_v1.1 (Illumina cat. #20015239). H9 cells were grown on Matrigel (Corning cat. #354234)-coated plates in mTesR1 medium (Stem cell technologies cat. #05851). We used EDTA (home-made) for splitting. For all nucleofections, the Human Stem Cell Nucleofector Kit 1 (Amaxa cat. #VVPH-5012) was used. For overexpression experiments, constructs containing CAG Promotor, cDNA of interest followed by p2A-eGFP were co-transfected with Sleeping beauty (Ivics et al., 1997) transposase to generate stable clones. Cells were detached with Accutase (Sigma-Aldrich cat. #A6964). We used 1 × 10⁶ cells per nucleofection. After nucleofection, cells were seeded into Matrigel (Corning cat. #354234)-coated 24-well plates in mTesR1 (Stem cell technologies cat. #05851) containing 10 µM Y-27632 (Calbiochem cat. #688000). Cells were further expanded into 6-well plates and sorted by FACS.

METHODS DETAIL

Generation of the *Htatsf1* flox allele

The *Htatsf1*^{flox} allele was generated by homologous recombination in mESCs. The targeting vector was obtained by sequential insertion of the following sequences in the 5' to 3' direction into the *HTATSF1* BAC by recombination-mediated genetic engineering in *Escherichia coli*; a fragment containing a neomycin cassette flanked with loxP sites was inserted downstream of *HTATSF1* Exon6. After selection Cre recombination was induced to remove the neomycin cassette. In a second step, a fragment containing a neomycin-resistance gene (flanked by *FRT* sites) and a loxP site was inserted upstream of the third exon of *HTATSF1*. In a third step, the targeting vector was generated by excision and insertion of the integrated sequences together with the flanking 5' and 3' homology arms by recombination-mediated genetic engineering from the modified BAC into the pBV-DTA-pA plasmid containing a gene encoding the herpes simplex virus thymidine kinase (for negative selection). DNA was linearized with PstI and transfected by electroporation into cells (1 × 10⁷) of the hybrid C57/BL6 × 129/Sv ES cell line A9, followed by selection with 250 µg/ml G418 (Invivogen cat. #ant-gn-1). PCR-positive clones were verified by Southern blot analysis before injection into C57/BL6J host blastocysts. Chimeras were identified based on coat color and mice carrying the *Htatsf1*^{flox} allele were backcrossed to the C57/BL6J strain. Mice were crossed to *Gt(ROSA)26Sor*^{tm1(CreERT2)Tyj}/J (Jackson Laboratories cat. #008463) (C57/BL6J background) to generate *Htatsf1*^{f/f}; CreERT2^{+/−} blastocysts for ESC isolation. Genotyping of mice was performed using the following oligonucleotides: CreERT2_for: 5'-CTGCCACGACCAAGTGACAGCAATG-3'; CreERT2_rev: 5'-GCCTTCTCTACACCTGCGGTGCT-3'; *HTATSF1_flox_1*: 5'-CTGGCAATGCCACTGACTGTAC-3'; *HTATSF1_flox_2*: 5'-AAAAGCCGTAGTTAGCCTGA-3'; *HTATSF1_flox_3*: 5'-AACATGTGGATCCACCTAATAACTTC-3'.

Mouse and human embryoid body generation

Mouse embryoid body generation and differentiation was performed as described previously with slight modifications (Kamiya et al., 2011). Briefly, mouse ESCs were detached using Trypsin-EDTA (GIBCO cat. #25300054). 4000 cells were seeded into the wells of an ultra-low-cluster 96-well plate (Szabo-Scandic cat. #COR7007) using Differentiation Medium containing GMEM (Sigma-Aldrich cat. #G5154), 10% KOSR (GIBCO cat. #10828028), Glutamax (Thermo Fisher Scientific cat. #35050-038), 1% Sodium Pyruvate (Sigma-Aldrich cat. #P2256), 1% MEM-NEAA (Sigma-Aldrich cat. #M7145) and 0.1 mM β-Mercaptoethanol (Sigma-Aldrich cat. #60-24-2). Medium was changed every other day. From day 2 on, EBs were moved from 96 wells to 6 cm dishes and incubated on an orbital shaker placed in the incubator.

Human EBs were generated from H9 ESCs grown in mTeSR1 (Stem cell technologies cat. #05851) as previously described (Lancaster et al., 2013). Briefly, human H9 embryonic stem cells were cultivated on Matrigel (Corning cat. #354234)-coated dishes in

mTeSR1 (Stem cell Technologies cat. #05851) medium. For generation of embryoid bodies (EBs), cells were split using Accutase (Sigma-Aldrich cat. #A6964) and seeded in an ultra-low-cluster 96-well plate (Szabo-Scandic cat. #COR7007) using 9000 cells per well in a volume of 150 µL using hES medium (DMEM/F12 (GIBCO cat. #11320-074), 20% KOSR (GIBCO cat. #10828028), 3% FBS (GIBCO cat. #10270-106), 1:100 Glutamax (Thermo Fisher Scientific cat. #35050-038), 1:100 MEM-NEAA (Sigma-Aldrich cat. #M7145), β-Mercaptoethanol (Sigma-Aldrich cat. #60-24-2)) with 4ng/mL bFGF (in house) and 50uM Y-27632 (Calbiochem cat. #688000).

On day 3, medium was changed to hES medium only. On day 6, EBs were transferred to Costar Ultra-low cluster 24-well plates (Szabo-Scandic cat. #COS3473) in neural induction media containing DMEM/F12 (GIBCO cat. #11320-074), 1:100 N2-supplement (Thermo Fisher Scientific cat. #17502001), Glutamax (Thermo Fisher Scientific cat. #35050-038), MEM-NEAA (Sigma-Aldrich cat. #M7145), and 1ug/ml Heparin sodium salt (Sigma-Aldrich cat. #H3149-10KU). EBs were fed again on day 8. Material was harvested on day 1, day 3, day 7 and day 10 of EB formation.

EpiSC generation from mouse ESCs

Differentiation of mouse ESCs to Epiblast stem cells (EpiSCs) was performed as described previously with slight modifications ([Tosolini and Jouneau, 2015](#)). Briefly, cells were seeded into Fibronectin (Sigma-Aldrich cat. #F0895)-coated 6-well plates at a density of $14,25 \times 10^4$ cells per 6 well in 2i medium. After 24 hours, the medium was changed to N2B27 medium containing DMEM/F12 (GIBCO cat. #11320-074), Neurobasal medium (GIBCO cat. #21103-049), 1% KOSR (GIBCO cat. #10828028), 5.25% BSA (VWR International cat. #422351S), Glutamax (Thermo Fisher Scientific cat. #35050-038), 1% MEM-NEAA (Sigma-Aldrich cat. #M7145), N2-supplement (Thermo Fisher Scientific cat. #17502001), B27-supplement (Thermo Fisher Scientific cat. #17504044) and 0.001 mM β-Mercaptoethanol (Sigma-Aldrich cat. #60-24-2). 20 ng/mL ActivinA (R&D Systems cat. #338-AC-050) and 12 ng/mL FGF-2 (home-made) were added freshly.

Cells underwent a highly selective process of differentiation with many cells dying in the first days. After several days, small colonies started to emerge. EpiSC colonies were split using Collagenase (GIBCO cat. #17104019). Cells were passaged 1 – 2 times on Fibronectin and then moved to MEFs and further grown on MEFs. EpiSCs used for experiments were used after passaging at least 2-3 times on MEFs.

Reversion of EpiSC to rESCs was performed as described previously ([Bao et al., 2009](#)). Briefly EpiSC colonies were treated with collagenase for 8min at room temperature. EpiSC colonies were removed from the feeder layer and cut into small pieces. Pieces were transferred to dishes with MEFs and standard ES cell medium (DMEM (GIBCO cat. #11960-044) containing 15% FBS (GIBCO cat. #10270-106), Glutamax (Thermo Fisher Scientific cat. #35050-038), 1% Sodium Pyruvate (Sigma-Aldrich cat. #P2256), 1% MEM-NEAA (Sigma-Aldrich cat. #M7145), 0.1 mM β-Mercaptoethanol (Sigma-Aldrich cat. #60-24-2) and 10^3 U/mL ESGRO mouse LIF (Millipore cat. #ESG1107)). After 15- 20 days cell clusters resembling ESC colonies emerged. rESC colonies were split using Trypsin-EDTA (GIBCO cat. #25300054) and cultured on MEFs.

Colony reforming efficiency assay

Htatsf1^{fl/fl}; CreERT2^{+/-} mouse ESCs were cultured in 2i/LIF as described. 4-OH-Tamoxifen (Sigma-Aldrich cat. #H7904) or Ethanol was added to cultures for 72h. Cells were split, and 1000 cells were plated per well of a 12 well plate. After 48h colonies were imaged using Celigo High Throughput Image Cytometer (Nexcelom). Colonies were detected and counted using Celigo Software.

Generation of naive human ES cells from human ES cells

Reversion of human ES cells to naive human ES cells was performed as previously described ([Takashima et al., 2014](#)). Briefly, H9 cells were nucleofected with Sleeping Beauty transposase ([Ivics et al., 1997](#)) constructs (CAG-hNanog-p2A-eGFP; CAG-hKLF2-p2A-eGFP; CAG-HTATSF1-p2A-eGFP). Cells were FACS sorted and expression of transgenes was confirmed by PCR. Cells were seeded on MEFs in medium containing Neurobasal medium (GIBCO cat. #21103-049), DMEM/F12 (GIBCO cat. #11320-074), N2-supplement (Thermo Fisher Scientific cat. #17502001), B27-supplement (Thermo Fisher Scientific cat. #17504044), 0.05% BSA, Glutamax (Thermo Fisher Scientific cat. #35050-038), 1% MEM-NEAA (Sigma-Aldrich cat. #M7145), 0.1 mM β-Mercaptoethanol (Sigma-Aldrich cat. #60-24-2), supplemented with 1000 U/mL human LIF (Peprotech cat. #AF-300-05), 1 µM PD032901 (Calbiochem cat. #444968) and 1 µM CHIR99021 (Calbiochem cat. #361571). Medium was changed daily. After 2 weeks, 2 µM Gö6983 (Sigma-Aldrich cat. #G1918) was added to the medium. AP staining was performed using Vector Blue AP substrate kit (Vector Laboratories cat. #SK-5300) according to the manufacturer's instructions.

Generation of synthetic embryos (ETS embryos)

ETS embryo generation was performed as described ([Harrison et al., 2017](#)). Briefly, ESCs were washed with PBS and treated with 0.05% Trypsin-EDTA (GIBCO cat. #25300054) for 2-3 minutes at 37°C. 2i medium (Neurobasal medium (GIBCO cat. #21103-049), DMEM/F12 (GIBCO cat. #11320-074), N2-supplement (Thermo Fisher Scientific cat. #17502001), B27-supplement (Thermo Fisher Scientific cat. #17504044), 2% FBS (GIBCO cat. #10270-106), 0.05% BSA, Glutamax (Thermo Fisher Scientific cat. #35050-038), 1% MEM-NEAA (Sigma-Aldrich cat. #M7145), 0.1 mM β-Mercaptoethanol (Sigma-Aldrich cat. #60-24-2)) was added and cells were centrifuged for 5 minutes at 1000 rpm. Cells were resuspended in 2i medium and centrifuged again. Finally, cells were resuspended in PBS.

The trophoblast stem cell line used stably and ubiquitously expressed GFP (Tanaka et al., 1998). TSCs were washed with PBS and treated with 0.05% Trypsin-EDTA (GIBCO cat. #25300054) for 3-4 minutes at 37°C, then cells were resuspended in TSC medium (RPMI 1640 medium (Sigma cat. #M3817) with 20% FBS (GIBCO cat. #10270-106), 2mM L-glutamine (GIBCO cat. #25030), 0.1mM β-Mercaptoethanol (Sigma-Aldrich cat. #60-24-2), 1mM sodium pyruvate (Sigma-Aldrich cat. #P2256), FGF4 (Peprotech cat. #100-31) and heparin (Sigma-Aldrich cat. #H3149-10KU)), centrifuged for 5 minutes at 1000 rpm and resuspended in TSC medium.

For both cell suspensions 7000 – 10000 cells were counted and mixed in a 1:1 ratio. This mixture was centrifuged, and cells resuspended in cold liquid Matrigel (Corning cat. #354234) (7500 cells per 25μl of Matrigel). Drops of Matrigel-cell suspension were plated and allowed to solidify. Once solid, medium consisting of 50% RPMI 1640 medium (Sigma cat. #M3817), 25% DMEM/F-12 (GIBCO cat. #11320-074) and 25% Neurobasal medium (GIBCO cat. #21103-049) supplemented with 10% FBS (GIBCO cat. #10270-106), 2mM L-glutamine (GIBCO cat. #25030), 0.1mM β-Mercaptoethanol (Sigma-Aldrich cat. #60-24-2), 0.5mM sodium pyruvate (Sigma-Aldrich cat. #P2256), 0.25x N2-supplement (Thermo Fisher Scientific cat. #17502001), 0.5x B27-supplement (Thermo Fisher Scientific cat. #17504044), 12.5 ng/ml FGF4 (Peprotech cat. #100-31) and 500ng/ml heparin (Sigma-Aldrich cat. #H3149-10KU) or ETS-embryo medium (Cell guidance systems cat. #M13) was added, and cells cultured at 5% CO₂ and 37°C. Every other day ETS embryos were fed.

ChIP-qPCR and ChIP-sequencing

Chromatin-IP for HTATSF1: Antibody against HTATSF1 was produced in rabbit against mouse HTATSF1 peptide SGNNLSGNDEFDEQLR-C coupled to KLH. Cells were grown to 80% confluence on 15 cm dishes. For every 10 mL of medium 1,1 mL formaldehyde solution (11% formaldehyde, 50mM HEPES pH 7.3, 100mM EDTA pH 8.0, 0.5mM EGTA pH8.0) were added to the cell culture dish and incubated at room temperature for 15 minutes. For every 10ml liquid in the dish 500μl 2.5M glycine were added. Medium was removed, and cells rinsed once with PBS. Cells were collected by scraping them off in PBS and washed twice in PBS. The pellet was re-suspended in 5 mL lysis buffer 1 (LB1) (50mM HEPES-KOH pH 7.5, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% Igepal, 0.25% Triton X-100) and incubated for 10 minutes at 4°C on a roller. After centrifugation, the pellet was re-suspended in 5 mL LB2 (10mM Tris-HCl pH 8.0, 200mM NaCl, 1mM EDTA pH8.0, 0.5mM EGTA pH8.0) and incubated as before. Cells were re-suspended in sonication buffer (50mM Tris-HCl pH 7.4, 140mM NaCl, 1mM EDTA, 1mM EGTA, 1% Triton X-100, 0.1% Na-deoxycholate (Thermo Fisher Scientific cat. #89904), 0.1% SDS) and sonicated on ice using a microtip (8 cycles for 30 s each, with 60 s breaks in between cycles at 18 W). After sonication 300μl of 10% Triton X-100 were added and mixed by pipetting. Samples were distributed on 2 1.5 mL tubes and centrifuged in a table top centrifuge at 13k rpm for 10 minutes at 4°C. 100μl of magnetic Dynabeads protein A (Thermo Fisher Scientific cat. #10002D) were washed 3 times with blocking solution (BS) (PBS containing 0.5% BSA) and re-suspended in BS. Antibody was added and pre-incubated overnight at 4°C on a roller. Beads were washed 3 times with BS and re-suspended in 100μl BS. The prepared beads and antibody were then mixed with the sonicated samples and incubated overnight at 4°C on a roller. Supernatant was removed using a magnetic tube rack and beads washed 3 times with sonication buffer, once with sonication buffer + 500mM NaCl, once with LiCl wash buffer (20mM Tris HCl pH 8.0, 1mM EDTA, 250mM LiCl, 0.5% NP-40 substitute (Sigma-Aldrich cat. #74385), 0.5% Na-deoxycholate (Thermo Fisher Scientific cat. #89904)), and once with TE. Beads were re-suspended in 210μl elution buffer (50mM Tris HCl pH 8.0, 10mM EDTA pH 8.0, 1% SDS) and incubated at 65°C for 15 minutes. During the incubation samples were vortexed every 2 minutes. Finally, samples were centrifuged at 13000 rpm for 1 minute at room temperature, 200μl of the supernatant transferred into a new tube, and incubated at 65°C overnight to reverse crosslinking. 1 volume of TE and RNase A (final concentration 0.5mg/ml) were added and incubated at 37°C for 30 minutes while shaking. Then CaCl₂ (final concentration 10mM) and proteinase K (final concentration 0.5mg/ml) were added and incubated at 55°C for 1 hour while shaking. 1 volume of phenol-chloroform was added, vortexed for 1 minute and centrifuged at 13000 rpm in a table top centrifuge for 10 minutes. 1/10 volume of NaAcetate (pH 5.2), 1μl GlycoBlue (Ambion cat. #AM9515) and 2.5 sample volumes of EtOH were added, mixed, and put to -80°C for 30 minutes. Samples were centrifuged at 13000 rpm for 30 minutes at 4°C. Pellets were washed once with 70% EtOH, dried, re-suspended in DNase free water, and used for qPCR (ChIP-qPCR) or submitted for sequencing (ChIP-Seq).

Chromatin-IP for Pol II was performed as described (Rahl et al., 2010). Briefly, for every 10 mL of medium 1,1 mL formaldehyde solution (11% formaldehyde, 50mM HEPES pH 7.3, 100mM EDTA pH 8.0, 0.5mM EGTA pH8.0) were added to the cell culture dish and incubated at room temperature for 15 minutes and washed twice with PBS. 100 μL of magnetic Dynabeads protein A (Thermo Fisher Scientific cat. #10002D) were blocked with 0.5% BSA (w/v) in PBS. Magnetic beads were bound with 10 μg of PolII antibody (Santa Cruz cat. #sc-899). Cells were scraped and lysed with lysis buffer 1 (50 mM HEPES pH 7.3, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40 substitute (Sigma-Aldrich cat. #74385), and 0.25% Triton X-100) and washed with lysis buffer 2 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA pH 8.0 and 0.5 mM EGTA pH 8.0). Cells were then resuspended and sonicated in sonication buffer (50 mM Tris-HCl pH 7.5, 140 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100, 0.1% Na-deoxycholate (Thermo Fisher Scientific cat. #89904), 0.1% SDS) using a microtip (8 cycles for 30 s each, with 60 s breaks in between cycles at 18 W). Sonicated lysates were incubated overnight at 4°C with beads bound with antibody. Beads were washed three times with sonication buffer, one time with sonication buffer with 500 mM NaCl, one time with LiCl wash buffer (20 mM Tris pH 8.0, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40 substitute (Sigma-Aldrich cat. #74385), 0.5% Na-deoxycholate (Thermo Fisher Scientific cat. #89904)), and one time with TE. Beads were re-suspended in 210μl elution buffer (50mM Tris HCl pH 8.0, 10mM EDTA pH 8.0, 1% SDS) and incubated at 65°C for 15 minutes. During the incubation samples were vortexed every 2 minutes. Finally, samples were centrifuged at 13000 rpm for 1 minute at room temperature, 200μl of the supernatant transferred into a new tube, and incubated

at 65°C overnight to reverse crosslinking. 1 volume of TE and RNase A (final concentration 0.5mg/ml) were added and incubated at 37°C for 30 minutes while shaking. Then CaCl₂ (final concentration 10mM) and proteinase K (final concentration 0.5mg/ml) were added and incubated at 55°C for 1 hour while shaking. 1 volume of phenol-chloroform was added, vortexed for 1 minute and centrifuged at 13000 rpm in a table top centrifuge for 10 minutes. 1/10 volume of NaAcetate (pH 5.2), 1µl GlycoBlue (Ambion cat. #AM9515) and 2.5 sample volumes of EtOH were added, mixed, and put to -80°C for 30 minutes. Samples were centrifuged at 13000 rpm for 30 minutes at 4°C. Pellets were washed once with 70% EtOH, dried, re-suspended in DNase free water and submitted for sequencing (ChIP-Seq).

ChIP-Sequencing: DNA from ChIP was used for library preparation. Libraries were generated using Kapa HTP Library Preparation Kit (Roche cat. #07961901001) according to the manufacturers' instructions. Barcoded samples were multiplexed and sequenced 50bp SE on a HighSeq2500 (Illumina). Library preparation and sequencing was performed at the VBCF NGS Unit (<https://www.vbcf.ac.at/>).

Western blots and co-IP

Preparation of nuclear extracts: All buffers contained Complete protease inhibitor cocktail (Roche cat. #05 056489001) and Phospho-Stop (Roche cat. #04 906837001). Cells with or without UV-crosslink were suspended in lysis buffer (10mM Tris HCl pH 7.4, 10mM NaCl, 15mM MgCl₂, 250mM Sucrose, 0.5% NP-40 substitute (Sigma-Aldrich cat. #74385)), vortexed for 10 s and kept on ice for 15 minutes. Samples were then centrifuged through a sucrose cushion (30% sucrose, 10mM Tris HCl pH 7.4, 10mM NaCl, 3mM MgCl₂) at 1300 g for 10 minutes at 4°C. The pellet was washed with cold 10mM Tris HCl pH 7.4 containing 10mM NaCl. The isolated nuclei were suspended in a buffer containing 10mM Tris HCl pH 7.4, 150mM NaCl, and 0.5% NP-40 substitute (Sigma-Aldrich cat. #74385) and sonicated for 30 s. Afterward samples were kept on ice for 30 minutes and then centrifuged in a table top centrifuge at 13000 rpm for 10 minutes at 4°C. The supernatant (crude nuclear extract) was stored at -80°C or the protein concentration determined with the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific cat. #23227) following the manufacturers protocol.

IP: All buffers contained Complete protease inhibitor cocktail (Roche cat. #05 056489001) and PhosphoStop (Roche cat. #04 906837001). Beads were washed with IP buffer (20mM Tris HCl pH 7.4, 150mM NaCl, 1mM Na₂EDTA, 1mM EGTA, 1% NP-40 substitute (Sigma-Aldrich cat. #74385), 1% Na-deoxycholate (Thermo Fisher Scientific cat. #89904)) and preincubated with antibody in IP buffer for 1 hour at 4°C. For a control the antibody was also pre-incubated with blocking peptide in IP buffer for 1 hour at 4°C. The nuclear extract samples (500 µg) were then added to the preincubated beads, the total volume increased to 500µl with IP buffer and incubated overnight at 4°C. Samples were then washed 5 times with IP buffer, re-suspended in Laemmli buffer and used for western blotting.

Western Blot: Samples were mixed with Laemmli buffer, loaded on precast gels (Invitrogen NuPage 4%-12% Bis-Tris Gel cat. #NP0321BOX) and run at 120 Volts until desired. Proteins were then transferred to a nitro-cellulose membrane using a wet blotting apparatus at 100 Volts for 1 hour. Membranes were then blocked with blocking solution (PBS containing 5% skimmed milk powder and 0.2% Tween 20) for at least 30 min and then incubated with primary antibody (diluted 1:1000 in blocking solution) overnight at 4°C. Then samples were washed 3 times with PBS containing 0.2% Tween 20 and incubated with secondary HRP-labeled antibody (diluted 1:2000 in blocking solution) for 1 hour at room temperature. After washing again 3 times with PBS containing 0.2% Tween 20 membranes were analyzed using western blot development kit (GE Healthcare cat. #RPN 2106) according to the manufacturers protocol.

For quantitative Western Blots the protocol was modified as follows:

Membranes used: Odyssey Nitrocellulose Membrane (LI-COR Biosciences cat. #926-31092)

Secondary antibodies: IRDye 680RD goat anti-rabbit (LI-COR Biosciences cat. #926-68071)

IRDye 680RD goat anti-mouse (LI-COR Biosciences cat. #926-68070)

IRDye 800CW goat anti-rabbit (LI-COR Biosciences cat. #926-32211)

IRDye 800CW goat anti-mouse (LI-COR Biosciences cat. #926-32210)

IRDye 800CW donkey anti-goat (LI-COR Biosciences cat. #926-32214)

After incubation with secondary antibodies, membranes were washed one additional time in PBS and air-dried. Finally, membranes were scanned using a LI-COR Odyssey CLx device. Analysis of band intensity was done using Fiji ([Schindelin et al., 2012](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3407733/)).

Protein synthesis measurements using OP-Puromycin Incorporation

Protein synthesis measurements in ESCs: To measure protein synthesis, mouse or human ESCs cells were incubated for 30 min in medium supplemented with O-propargyl-puromycin (OPP) (Jenabioscience cat. #NV-931-05) (20 µg/ml final concentration). Cells were then harvested, washed with PBS, fixed with 1% PFA for 15 min on ice, and then permeabilized using PBS supplemented with 3% FBS (GIBCO cat. #10270-106) and 0.1% saponin for 5 min at room temperature. The copper-catalyzed azide-alkyne cyclo-addition was done using the Click-iT Plus OPP Alexa Fluor 488 or 647 Protein Synthesis Assay Kit (ThermoFisher Scientific cat. #C10456 or cat. #C10458, respectively) according to the manufacturer's protocol. Cells were re-suspended in 200 µL PBS supplemented with 3% FBS (GIBCO cat. #10270-106) and 0.1% saponin and analyzed by flow cytometry.

Protein synthesis measurements in mouse embryos: Female mice were paired with males in the evening and on the following morning examined for vaginal plugs. Females with positive vaginal plugs were considered as 0.5 days post-coitum (dpc) or

embryonic day 0.5 (E0.5) and embryos were staged accordingly. E4.5 embryos were flushed from the uterus with M2 medium (Sigma-Aldrich cat. #M7167). E5.5-E7.5 embryos were dissected in M2 medium (Sigma-Aldrich cat. #M7167). Mouse embryos (E4.5-E7.5) were incubated for 60 minutes in EmbryoMax KSOM medium (Millipore cat. #MR-020P-5F) supplemented with O-propargyl-puro-mycin (OPP) (Jenabioscience cat. #NV-931-05) (20 µg/ml final concentration). Embryos were fixed with 4% PFA for 10-20 min (depending on the age of the embryo) at 37°C and permeabilized in PBS supplemented with 0.5% Triton X-100 for 20 min at room temperature and blocked with blocking solution (PBS supplemented with 0.05% Triton X-100, 5% FBS (GIBCO cat. #10270-106), 5% BSA (VWR International cat. #422351S), and 0.1% sodium-azide). The copper-catalyzed azide-alkyne cycloaddition was done using the Click-iT Plus OPP Alexa Fluor 488 Protein Synthesis Assay Kit (ThermoFisher Scientific cat. #C10456) according to the manufacturer's protocol. Afterward, embryos were co-stained with antibodies (see Immunohistochemistry). Epiblast was identified by co-staining with OCT4. Images were acquired on a confocal microscope (Zeiss 780). Where indicated, composite confocal images were taken using the tiles Function in the Zeiss Zen software (Zeiss). Nuclei were identified manually and mean gray values of selected areas were measured semi-automatically using Fiji ([Schindelin et al., 2012](#)).

Isolation of Nucleoli

Nucleoli isolation was performed as described ([Li and Lam, 2014](#)). Cells were either UV crosslinked or left without crosslink. Medium was removed from dishes and cells washed twice with cold solution I (0.5M sucrose, 3mM MgCl₂, Complete Protease Inhibitor Cocktail (Roche cat. #05056489001)). Cells were scraped off, transferred to a tube, solution I added to a volume of 0.5ml and sonicated. The cell suspension was underlaid with 0.7ml solution II (1.0M sucrose, 3mM mgCl₂, Complete Protease Inhibitor Cocktail (Roche cat. #05056489001)) and centrifuged at 1800 g for 5 minutes at 4°C. Sucrose fractionation was performed 2 times to enhance purity of Nucleoli. The pellet containing the nucleoli was resuspended in a buffer suitable for downstream applications (see RNA-IP or Western Blots).

RNA-IP

1x10⁷ cells were suspended in 2 mL freshly prepared nuclear isolation buffer (1.28M sucrose, 40mM Tris HCl pH 7.4, 20mM MgCl₂, 4% Triton X-100) and 6 mL water and kept on ice for 20 minutes with frequent mixing. Cells were centrifuged at 2500 g for 15 minutes and re-suspended in 1 mL freshly prepared RIP buffer (150mM KCl, 25mM Tris pH 7.4, 5mM EDTA, 0.5mM DTT, 0.5% NP-40 substitute (Sigma-Aldrich cat. #74385), 100U/ml RNase inhibitor SUPERaseIn (Ambion cat. #AM2694), Complete Protease Inhibitor Cocktail (Roche cat. #05056489001)). Alternatively, isolated nucleoli were resuspended in RIP buffer. Chromatin was sheared mechanically with 15-20 strokes of a dounce homogenizer and cell debris removed by centrifugation at 13000 g for 10 minutes. 5µg of antibody were added to the supernatant and incubated at 4°C overnight rotating. Rmp protein A Sepharose fast flow (GE Healthcare cat. #17-5138-01) beads (40µl) were added and incubated for 1 hour at 4°C with rotation. Beads were pelleted at 2.500 rpm for 30 s and washed 3 times with 500µl RIP buffer and once with PBS. To isolate co-precipitated RNAs, beads were re-suspended in Trizol (Thermo Fisher Scientific cat. # 15596018) following the manufacturers protocol. Finally, RNA was eluted with nuclease free water.

RNA-sequencing

RNA was isolated using Trizol (Thermo Fisher Scientific cat. #15596018) according to the manufacturer's instructions. RNA concentration and integrity were analyzed using RNA 6000 Nano Chip (Agilent Technologies cat. #5067-1511). RNA was enriched for mRNA using Dynabeads mRNA Purification Kit (Thermo Fisher Scientific cat. # 61006). RNA was fragmented, and Reverse Transcription was performed using SuperScript III Reverse Transcriptase (Thermo Fisher Scientific cat. #18080044). Actinomycin D (Sigma-Aldrich cat. #A1410) was added to ensure maximal strand specificity. Second strand synthesis was performed, and libraries were generated using Kapa HTP Library Preparation Kit (Roche cat. #07961901001). Barcoded samples were multiplexed and sequenced 100bp PE on a HighSeq2000 (Illumina). Library preparation and sequencing was performed at the VBCF NGS Unit ([www.vbcf.ac.at](#)).

RNA extraction and qRT-PCR

RNA was isolated using Trizol (Thermo Fisher Scientific cat. #15596018) according to the manufacturer's instructions. Samples were treated with Turbo DNA-Free Kit (ThermoFisher Scientific cat. #AM1907). cDNA generation was performed using SuperScript III Reverse Transcriptase (Thermo Fisher Scientific cat. #18080044) with random primers according to the manufacturer's instructions. qPCR analysis was performed on a CFX96/CFX384 Real-Time System/C1000 Thermal Cycler (BioRad).

Plasmids

For overexpression constructs, based on the Sleeping Beauty Transposase System, the pCMV promoter from pCMV(CAT)T7-SB100 (Addgene cat. #34879) ([Mátés et al., 2009](#)) was replaced with CAG Promotor from pCAGEN (Addgene cat. #11160) ([Matsuda and Cepko, 2004](#)). IRDR-R and IRDR-L sequences from pT2/LTR7-GFP (Addgene cat. #62541) were cloned into pCAGEN (Addgene cat. #11160) to produce pCAG-GS/IR. cDNAs used for overexpression were amplified from mouse or human cDNA and cloned into the MCS of pCAG-GS/IR. For Crispr/Cas9 mediated deletion of p53 and PTEN in mouse ESCs, we used pSpCas9n(BB)-2A-Puro (PX462) V2.0 (Addgene cat. #62987) or pSpCas9n(BB)-2A-GFP (PX461) (Addgene cat. #48140) ([Ran et al., 2013](#)).

Polysome Extraction and analysis

Polysome fractionation was performed as previously described with slight modifications (Aeschimann et al., 2015). RNAs were immobilized on ribosomes by adding medium containing 100 μ g/ml Cycloheximide (Sigma-Aldrich cat. #C1988) and incubating them for 10 minutes in an incubator. Cells were washed twice with PBS + 100 μ g/ml Cycloheximide (Sigma-Aldrich cat. #C1988) and scraped in 1ml of PBS/ Cycloheximide (Sigma-Aldrich cat. #C1988). Cells were centrifuged at 500 g for 5 minutes at 4°C and the pellet re-suspended in 600 μ l lysis buffer (20mM HEPES pH 7.4, 150mM KCL, 15mM MgCl₂, 100 μ g/ml Cycloheximide (Sigma-Aldrich cat. #C1988), 1mM DTT, 2% NP-40 substitute (Sigma-Aldrich cat. #74385), Complete protease inhibitor cocktail (Roche cat. #056489001), 10U/ml RNasin (Promega cat. #N2515)). After 15 minutes on ice cells were centrifuged at 13000 rpm for 10 minutes at 4°C. The supernatant was transferred to a fresh tube and loaded onto a 15% to 60% continuous sucrose gradient (15% and 60% sucrose solutions contained 15% (w/v) or 60% (w/v) sucrose, respectively and 20mM HEPES pH 7.4, 150mM MgCl₂, 100 μ g/ml Cycloheximide (Sigma-Aldrich cat. #C1988), 1mM DTT). Samples were then centrifuged in a Beckmann ultracentrifuge using an SW40 rotor at 39000 rpm for 3 hours at 4°C. Finally, fractions were separated and polysome distribution was analyzed.

Mass spectrometry

IP for MS

Mouse A9 ESCs were grown on MEFs with conventional ESC medium. Nuclear extracts were prepared as described above. All buffers contained Complete protease inhibitor cocktail (Roche cat. #056489001) and PhosphoStop (Roche cat. #04906837001). Beads were washed with IP buffer (20mM Tris HCl pH 7.4, 137mM NaCl, 2mM Na₂EDTA, 10% Glycerin, 1% Triton X-100) and pre-incubated with antibody in IP buffer for 1 hour at 4°C. For a control the antibody was also preincubated with blocking peptide in IP buffer for 1 hour at 4°C. The nuclear extract samples were then added to the pre-incubated beads, the total volume increased to 500 μ l with IP buffer and incubated overnight at 4°C. Samples were then washed 5 times with IP buffer, 2 times with high salt buffer (20mM Tris HCl pH 7.4, 300mM NaCl, 2mM Na₂EDTA, 10% Glycerin, 1% Triton X-100), 2x with NaCl solution (150 mM NaCl). Samples were loaded on precast NuPage 4%-12% Bis-Tris Gels (Invitrogen cat. #NP0321BOX) and stained with Pierce Silver Stain Kit (Thermo Fisher Scientific cat. #24600). Bands were cut, washed, reduced, alkylated and washed again. In-gel digest was performed using Trypsin-EDTA (GIBCO cat. #25300054) and peptides were extracted from gels and measured.

NanoLC-MS Analysis

The nano HPLC system used was an UltiMate 3000 HPLC RSLC nano system (Thermo Fisher Scientific, Bremen, Germany) coupled to an LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, Bremen, Germany), equipped with a Proxeon nanospray source (Proxeon, Odense, Denmark). Peptides were loaded onto a trap column (Thermo Fisher Scientific, Bremen, Germany, PepMap C18, 5 mm × 300 μ m ID, 5 μ m particles, 100 Å pore size) at a flow rate of 25 μ L min⁻¹ using 0.1% TFA as mobile phase. After 10 min, the trap column was switched in line with the analytical column (Thermo Fisher Scientific, Bremen, Germany, PepMap C18, 500 mm × 75 μ m ID, 3 μ m, 100 Å). Peptides were eluted using a flow rate of 230 nL min⁻¹, and either a binary 25 min gradient, respectively 70 min running time, or a 2h gradient, respectively 165 min.

Both gradients start with the mobile phases: 98% A (water/acetonitrile/formic acid, 97.4/2.5/0.1, v/v/v) and 2% B (water/acetonitrile/formic acid, 19.92/80/0.08, v/v/v). The 25 min gradient increases to 40% B over the next 25min, followed by a gradient in 5 min to 90% B, stays there for 5 min and decreases in 2 min back to the gradient 98% A and 2% B for equilibration at 30°C. The 2h gradient increases to 35% B over the next 120min, followed by a gradient in 5 min to 90% B, stays there for 5 min and decreases in 5 min back to the gradient 98% A and 2% B for equilibration at 30°C.

The LTQ Orbitrap Velos was operated in data-dependent mode, using a full scan in the Orbitrap (m/z range 350-2000, nominal resolution of 60 000, target value 1E⁶) followed by MS/ MS scans of the 12 most abundant ions in the linear ion trap. MS/MS spectra (normalized collision energy 35%; activation value q 0.25; activation time 10 ms; isolation width 2, target value 1E⁴) were acquired and subsequent activation was performed on fragment ions through multistage activation. The neutral loss mass list was therefore set to -98, -49, and -32.6 m/z. Precursor ions selected for fragmentation (charge state 2 and higher) were put on a dynamic exclusion list for 180 s. Additionally, singly-charged parent ions were excluded from selection for MS/MS experiments and the monoisotopic precursor selection feature was enabled.

Immunohistochemistry and Immunocytochemistry

Medium was removed, EBs were washed in PBS and fixed using pre-warmed 4% PFA at 37°C for 15 minutes. EBs were then embedded in PBS containing 7.5% gelatin (Sigma-Aldrich cat. #G1890) and 10% Sucrose. Solid gelatin blocks were then frozen for 1 minute in 2-Methylbutane (Sigma-Aldrich cat. #M32631) cooled to -40°C with dry ice. Blocks were stored at -80°C. Sections of 20 μ m thickness were prepared using a cryostat and put onto glass slides, which were then stored at -20°C.

Slides from -20°C were dried for 30 minutes at room temperature and washed in PBS once for 10 minutes on a shaker. Afterward tissue was permeabilized with 0.5% Triton X-100 for 20 minutes at room temperature and blocked in blocking solution (BS) (PBS containing 0.05% Triton X-100, 5% FBS (GIBCO cat. #10270-106), 5% BSA (VWR International cat. #422351S), and 0.1% sodium-azide) for 30 minutes at room temperature. Slides were then incubated with primary antibodies diluted in BS (1:200, except for α -htat (gp) 1:100) at 4°C overnight. Samples were washed in PBS 3 times for at least 5 minutes on a shaker and then incubated with fluorescent labeled secondary antibodies (1:800) and Hoechst (Thermo Fisher Scientific cat. #H-3569) (1:5000) in BS for 1 hour at room temperature in a light protecting container. Slides were washed in PBS 3 times for at least 5 minutes on a shaker, then briefly

washed with water and mounted in Fluorescent Mounting Medium (Dako cat. #S3023). Slides were finally dried at room temperature overnight and analyzed using a confocal microscope (Zeiss LSM780).

Cells were grown on glass coverslips coated with gelatin (Sigma-Aldrich cat. #G1890) or Matrigel (Corning cat. #354234). Medium was removed, cells were washed in PBS and then fixed in 4% PFA for 10 minutes at 37°C. Permeabilization was done with 0.5% Triton X-100 at room temperature for 20 minutes. Cells were blocked in blocking solution (BS) (PBS containing 0.05% Triton X-100, 5% FBS (GIBCO cat. #10270-106), 5% BSA (VWR International cat. #422351S), and 0.1% sodium-azide) for 30 minutes at room temperature and incubated with primary antibodies diluted in BS at 4°C overnight. Samples were washed in PBS 3 times for at least 5 minutes on a shaker and then incubated with fluorescent labeled secondary antibodies (1:800) and Hoechst (Thermo Fisher Scientific cat. #H-3569) (1:5000) in BS for 1 hour at room temperature in a light protecting container. Slides were washed in PBS 3 times for at least 5 minutes on a shaker, then briefly washed with water and mounted in Fluorescent Mounting Medium (Dako cat. #S3023). After drying overnight at room temperature cells were analyzed using a fluorescent microscope or on a confocal microscope (Zeiss LSM780).

Immunohistochemistry of mouse embryos and ETS embryos was performed as previously described (Harrison et al., 2017). Mouse embryos or ETS embryos were fixed in 4% PFA for 20 minutes at room temperature, washed twice in PBS containing 0.05% Tween-20 (PBT) and permeabilized with 0.3% Triton X-100 and 0.1% Glycin for 15 minutes at room temperature. Embryos were then blocked in blocking buffer (PBS plus 10% FBS (GIBCO cat. #10270-106) and 1% Tween-20). Incubation with primary antibodies was done overnight in blocking buffer at 4°C. After washing twice in PBT embryos were incubated in blocking buffer containing secondary antibodies and Hoechst (Thermo Fisher Scientific cat. #H-3569) overnight at 4°C. Finally, embryos were washed twice in PBT and mounted in Fluorescent Mounting Medium (Dako cat. #S3023).

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of ChIP-Sequencing data

Reads from ChIP-Seq were aligned with bowtie to the *Mus musculus* genome (mm9) (Langmead et al., 2009). Options are set to 3 mismatches, 1 max multihit, tryhard and best strata. Duplicates are removed with picard-tools (1.82) (<http://broadinstitute.github.io/picard>) BED files of the reads are generated with bedtools (Quinlan and Hall, 2010) and peaks are called with Macs 1.4 (Zhang et al., 2008).

Analysis for rDNA occupancy of HTATSF1: Reads of HTATSF1 ChIP and input were aligned against the *Mus musculus* ribosomal DNA, complete repeating unit (BK000964.3) with Bowtie (Langmead et al., 2009). The coverage was visualized in IGV. The relative track height was chosen according to the read count of the corresponding genomic alignments. Broad Peaks were called using MACS and fold enrichment and q Values were provided by the software (Zhang et al., 2008).

Analysis of Pol II traveling ratio: Deduplicated ChIP-seq alignments are converted to bigWigs and replicates are merged with the BigWig and BigBed tools (Kent et al., 2010). Tracks were normalized with RseQC to a wigsum of 200000000 (Wang et al., 2012). Promotor regions were defined from -30 to +300 and gene body is the remaining gene. Density of promotor and body was summarized with bigWigAverageOverBed and the traveling ratio was calculated for the average over bases with non-covered bases counting as zeroes for promotor over body (Rahl et al., 2010).

Analysis of RNA-sequencing data

The strand specific paired-end reads were screened for ribosomal RNA by aligning with BWA (v0.6.1) (Li and Durbin, 2009) against known rRNA sequences (RefSeq). The rRNA subtracted reads were aligned with TopHat (v1.4.1) (Trapnell et al., 2009) against the *Mus musculus* genome (mm9) and a maximum of 6 missmatches. Maximum multi-hits were set to 1 and InDels as well as Micro-exon-search was enabled. Additionally, a gene model was provided as GTF (UCSC, 2013_02_21, mm9).

Analysis of Intron Retention: The RefSeq annotation (UCSC, 2013_02_21, mm9) was prepared with DEXSeq (Bioconductor) and the intronic not covered regions are added as exonic parts (Anders et al., 2012). That annotation was used to count the aligned reads with dexpseq_count. Standard DEXseq analysis were performed on the counts for exons and introns. Dexseq counts are normalized by intron length. Introns with an adjusted p value < 0.01 and normalized count > 0.1 are selected. As a second approach we determined the intron coverage as the average over bases with non-covered bases counting as zeroes for normalized bigWigs (Kent et al., 2010). Coverage below 1 was set to 1. We calculated for each gene the ratio between the highest and second highest coverage. We applied a threshold of 0.5 to the fold change and intersect it with introns selected by DEXseq.

GO Term and KEGG pathway analysis was performed using DAVID Bioinformatics Analyses Tool (Huang et al., 2009a, 2009b).

Analysis of mass spectrometry data

For peptide identification, the RAW-files were loaded into Proteome Discoverer (version 1.3.0.339, Thermo Scientific). All hereby created MS/MS spectra were searched using Mascot 2.2.07 (Matrix Science, London, UK) against the NCBI non-redundant protein sequence database, using the taxonomy mouse (142,659 protein sequences). The following search parameters were used: Beta-methylthiolation on cysteine was set as a fixed modification, oxidation on methionine and substitution of glutamine against pyro-glutamic acid were set as variable modifications. Monoisotopic masses were searched within unrestricted protein masses for tryptic

enzymatic specificity. The peptide mass tolerance was set to ± 5 ppm and the fragment mass tolerance to ± 0.5 Da. The maximal number of missed cleavages was set to 3. The localization of modification sites within the peptides was performed with the tool phos- phoRS (Taus et al., 2011).

For better visualization, the results of the searches were loaded into Scaffold (version 3.3.1, Proteome Software Inc.), using a minimum of 2 unique peptides per protein and a Mascot Score of at least 20 as cut-off filters.

Statistics

Statistical analyses were performed with GraphPad Prism version 7 for Mac (GraphPad Software, La Jolla California, USA). Statistically significant threshold was accepted as $p < 0.05$. Statistical analysis of quantifications performed was done using unpaired Student's two-tailed t test for significance between two experimental groups in all experiments except NGS-based approaches and unless otherwise noted (see Figure Legends). Analysis of length of all introns and introns retained upon loss of HTATSF1: Mann-Whitney Test was used to evaluate significance.

Unless otherwise noted, n equals biological replicates from independent experiments. For quantification of immunohistochemistry in 1C-E, 2B, D and 6C-E n equals cells from at least 3 independent experiments. For quantification of immunohistochemistry in 7B, C n equals cells from 4 or 5 embryos. For quantification of colony reforming efficiency in 1G and 6G n equals wells from 4 independent experiments.

DATA AND SOFTWARE AVAILABILITY

The accession number for the RNA- and ChIP-Sequencing data reported in this paper is GEO: GSE96595.