Cross-Regulation between TDP-43 and Paraspeckles Promotes Pluripotency-Differentiation Transition

Graphical Abstract

Highlights

- TDP-43 maintains pluripotency by regulating expression of pluripotency factors
- TDP-43 represses formation of paraspeckles in ESCs by regulating Neat1
- The paraspeckle-inducing isoform of Neat1 promotes differentiation of ESCs and embryos
- Cross-regulation between TDP-43 and Neat1 enhances pluripotency-differentiation axis

Authors
Miha Modic, Markus Grosch, Gregor Rot, ..., Tetsuro Hirose, Jernej Ule, Micha Drukker

Correspondence
jernej.ule@crick.ac.uk (J.U.), micha.drukker@helmholtz-muenchen.de (M.D.)

In Brief
Modic et al. uncover opposing roles of TDP-43 and paraspeckles in pluripotency and differentiation that are further enhanced by their cross-regulation. TDP-43 represses paraspeckles through processing of the scaffolding lncRNA Neat1, whereas paraspeckles partially sequester TDP-43. This reciprocal relationship promotes coordinated changes in alternative polyadenylation essential for efficient exit from pluripotency.

Modic et al., 2019, Molecular Cell 74, 951–965
June 6, 2019 © 2019 The Authors. Published by Elsevier Inc.
https://doi.org/10.1016/j.molcel.2019.03.041
Cross-Regulation between TDP-43 and Paraspeckles Promotes Pluripotency-Differentiation Transition

Miha Modic,1,3,4 Markus Grosch,1 Gregor Rot,5 Silvia Schirge,1,6 Tjasa Lepko,1 Tomohiro Yamazaki,7 Flora C.Y. Lee,3,4 Ejona Rusha,1 Dmitry Shaposhnikov,1 Michael Palo,3,4 Juliane Meri-Pham,3 Davide Cacchiarelli,3,10,11 Boris Rogelj,12,13,14 Stefanie M. Hauck,6 Christian von Mering,9 Alexander Meissner,9,10,15 Heiko Lackert,1,6 Tetsuro Hirose,7 Jernej Ule,3,4,* and Micha Drukker12,16,*

1Institute of Stem Cell Research, Helmholtz Zentrum München, 85764 Neuherberg, Germany
2Comprehensive Pneumology Center (CPC-M), Ludwig-Maximilians-Universität München, Asklepios Fachkliniken München-Gauting und Helmholtz Zentrum München, Max-Lebsche-Platz 31, 81377 Munich, Germany
3The Francis Crick Institute, London NW1 1AT, UK
4Department for Neuromuscular Diseases, UCL Queen Square Institute of Neurology, London WC1N 3BG, UK
5Institute of Molecular Life Sciences of the University of Zurich and Swiss Institute of Bioinformatics, 8057 Zurich, Switzerland
6Institute of Diabetes and Regeneration Research, Helmholtz Zentrum München, 85764 Neuherberg, Germany
7Institute for Genetic Medicine, Hokkaido University, Sapporo 060-0815, Japan
8Research Unit Protein Science, Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, 80939 Munich, Germany
9Broad Institute of Harvard University/MIT, Cambridge, MA 02142, USA
10Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA
11Telethon Institute of Genetics and Medicine (TIGEM), NA 80078 Pozzuoli, Italy
12Department of Biotechnology, Jožef Stefan Institute, 1000 Ljubljana, Slovenia
13Faculty of Chemistry and Chemical Technology, University of Ljubljana, 1000 Ljubljana, Slovenia
14Biomedical Research Institute BRIS, 1000 Ljubljana, Slovenia
15Department of Genome Regulation, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany
16Lead Contact
*Correspondence: jernej.ule@crick.ac.uk (J.U.), micha.drukker@helmholtz-muenchen.de (M.D.)
https://doi.org/10.1016/j.molcel.2019.03.041

SUMMARY

RNA-binding proteins (RBPs) and long non-coding RNAs (lncRNAs) are key regulators of gene expression, but their joint functions in coordinating cell fate decisions are poorly understood. Here we show that the expression and activity of the RBP TDP-43 and the long isoform of the lncRNA Neat1, the scaffold of the nuclear compartment “paraspeckles,” are reciprocal in pluripotent and differentiated cells because of their cross-regulation. In pluripotent cells, TDP-43 represses the formation of paraspeckles by enhancing the polyadenylated short isoform of Neat1. TDP-43 also promotes pluripotency by regulating alternative polyadenylation of transcripts encoding pluripotency factors, including Sox2, which partially protects its 3' UTR from miR-21-mediated degradation. Conversely, paraspeckles sequester TDP-43 and other RBPs from mRNAs and promote exit from pluripotency and embryonic patterning in the mouse. We demonstrate that cross-regulation between TDP-43 and Neat1 is essential for their efficient regulation of a broad network of genes and, therefore, of pluripotency and differentiation.

INTRODUCTION

A long noncoding RNA (lncRNA) called Neat1 acts as a scaffold for paraspeckles by recruiting many RNA-binding proteins (RBPs) that have been implicated in development, cancer, and neurodegeneration, including TDP-43 and FUS (West et al., 2016). Paraspeckles have been implicated in post-transcriptional regulation by association with specific mRNAs and RBPs (Chen and Carmichael, 2009; Hennig et al., 2015; Jiang et al., 2017; Naganuma et al., 2012; Prasanth et al., 2005). Remarkably, paraspeckles have been identified in many types of somatic cells but not in embryonic stem cells (ESCs) (Chen and Carmichael, 2009). Several lncRNAs and RBPs can affect differentiation of ESCs by regulating gene expression (Flynn and Chang, 2014), but the role of their cross-regulation in promoting efficient transitions during differentiation is unknown. Thus, investigating the recruitment of specific RBPs by Neat1 into paraspeckles in the context of ESC differentiation can answer the larger question of how the scaffolding of RBPs by lncRNAs is coupled to cell fate transitions and how this might coordinate the broader gene regulatory networks that establish distinct cell identities.

Here we reveal the importance of cross-regulation between Neat1 and TDP-43 in the context of cellular differentiation. We find that an evolutionarily conserved switch in alternative polyadenylation (APA) of Neat1 is regulated by TDP-43 and leads to induction of the long isoform (Neat1_2), which scaffolds...
paraspeckles upon exit from pluripotency. We also show that TDP-43 and NEAT1.2 have opposing functions during differentiation because of their cross-regulation: TDP-43 represses the formation of paraspeckles in pluripotent cells, whereas NEAT1 partly sequesters TDP-43 away from mRNAs in differentiated cells. TDP-43 also globally regulates the APA of many mRNAs encoding pluripotency regulators, including the core pluripotency and reprogramming factor Sox2. We demonstrate that their cross-regulation is essential for the efficient functions of TDP-43 and Neat1 in promoting states of pluripotency and differentiation, respectively. This shows how a lncRNA can act together with cross-regulated RBPs to increase the efficiency of cell fate transitions.

RESULTS

APA Induces Formation of Paraspeckles upon Exit from Pluripotency

The NEAT1 gene produces two transcripts, a short isoform that is polyadenylated and does not form paraspeckles (NEAT1.1) and a full-length isoform that ends with a triple helix and forms paraspeckles that scaffold various RBPs (NEAT1.2) (Li et al., 2017; Naganuma et al., 2012). We quantified NEAT1 foci in human ESCs (hESCs) that were prompted to differentiate to diverse fates using single-molecule fluorescence in situ hybridization (FISH) probes that recognized either the region common to pluripotent cells (hESCs) that were prompted to differentiate to diverse fates and DAZAP1 (Fox et al., 2018), was decreased in NEAT1.2 hESCs (Figure 2D; Figure S2G, Table S2), demonstrating that induction of NEAT1.2 indeed leads to partial recruitment of a cohort of RBPs into paraspeckles and away from mRNAs. In addition, we found that TDP-43, but not control RBPs, was increasingly retained in the supernatant after depletion of polyadenylated mRNAs in NEAT1.2 hESCs (Figure 2C), confirming a role of NEAT1.2 in sequestering TDP-43 into paraspeckles.

TDP-43 is known to be localized to the outer shell of paraspeckles (Naganuma et al., 2012; West et al., 2016), but it has been unclear whether TDP-43 needs to bind to NEAT1.2 to localize to paraspeckles. The NEAT1.2 isoform contains three regions rich in UGUG repeats that overlap with the main peaks of TDP-43 binding (Tollervey et al., 2011). We used CRISPR-Cas9 editing in the HAP1 cell line, which forms paraspeckles in the steady state, to delete these three regions from the endogenous NEAT1 gene (Figure 2D; NEAT1.JUG) that are not essential for paraspeckle assembly (Yamazaki et al., 2018). We confirmed that NEAT1 foci formation and assembly of PSPC1 in paraspeckles remained intact in NEAT1.JUG cells (Figure 2E). In contrast, localization of TDP-43 into paraspeckles dramatically decreased in NEAT1.JUG cells (Figures 2F and 2G). To confirm that the UGUG-rich repeats in the NEAT1.2 isoform are essential for the localization of TDP-43 to paraspeckles, we inserted a stretch of 60 UG-repeats to the 3’ end of endogenous NEAT1.JUG (Figure 2H), which restored the colocalization of TDP-43 with paraspeckles (Figures 2I and 2J). Taken together, our findings indicate that TDP-43 localizes to paraspeckles because of its direct binding to the NEAT1.2 isoform, which whether this RBP regulates NEAT1. Indeed, we found that overexpression of TDP-43 repressed paraspeckle formation upon spontaneous differentiation of mESCs (Figure 1B; Figure S1O). Conversely, depletion of TDP-43 induced the production NEAT1.2 and paraspeckle formation in hESCs or mESCs to a similar extent as deletion of the pA site, as detected by single-molecule imaging and RNA-seq (Figures 1D–1F; Figure S2C). We conclude that the isoform switch of NEAT1 is regulated by TDP-43 to coincide with the exit from pluripotency, ensuring that NEAT1.2 scaffolds paraspeckles upon differentiation irrespective of the embryonic lineage.

NEAT1.2 Recruits TDP-43 into Paraspeckles and Away From mRNAs

Because TDP-43 is a known paraspeckle RBP (West et al., 2016), we next asked whether reciprocal regulation by NEAT1.2 on TDP-43 takes place upon early differentiation. We used a global mRNA-RBP occupancy analysis (Baltz et al., 2012; Castello et al., 2012) to examine how induction of NEAT1.2 in the undifferentiated state of NEAT1.pA hESCs affects the binding of RBPs to polyadenylated mRNAs (Figure 1D; Figures S2D and S2E). Because NEAT1.2 is not polyadenylated, it remains present in the supernatant after the removal of polyadenylated mRNAs (Figure S2F). Thus, the RBPs that are recruited to paraspeckles and crosslinked to NEAT1.2 in NEAT1.pA hESCs are not expected to co-purify with polyadenylated mRNAs, and, as a result, their mRNA occupancy is expected to decrease (Figure 2A). Indeed, the mRNA occupancy of known paraspeckle proteins, including TDP-43, NONO, SFQ, RBM14, HNRNPH3, HNRNPK, FUS, and DAZAP1 (Fox et al., 2018), was decreased in NEAT1.pA hESCs (Figure 2B; Figure S2G, Table S2), demonstrating that induction of NEAT1.2 indeed leads to partial recruitment of a cohort of RBPs into paraspeckles and away from mRNAs. In addition, we found that TDP-43, but not control RBPs, was increasingly retained in the supernatant after depletion of polyadenylated mRNAs in NEAT1.2 pA hESCs (Figure 2C), confirming a role of NEAT1.2 in sequestering TDP-43 into paraspeckles.

To gain insights into the mechanism that contributes to NEAT1 dynamics, we sequenced RNAs from undifferentiated hESCs and their progeny. Sequencing of nascent RNA-seq (nascent RNA-seq) by selecting for transcripts pulse-labeled with ethynyl uridine revealed that NEAT1.1 is expressed in undifferentiated hESCs, and NEAT1.2 was found among the 18 transcripts that were commonly upregulated within 24 h differentiation treatment by WNT3A, a small-molecule inhibitor of GSK3 (CHIR99021), and BMP4 (Figure 1C; Figure S1D; Table S1). As reported for somatic cells (Mao et al., 2011; Shertosov and Dundr, 2011), we found that NEAT1.2 is localized to the chromatin, whereas NEAT1.1 is localized to the nucleoplasm (Figures S1E–S1J). To determine whether the expression of NEAT1 is post-transcriptionally regulated, we used CRISPR-Cas9-mediated editing to delete the internal polyadenylation (pA) site of NEAT1 in hESCs and of Neat1 in mESCs. We found that this induced the production of NEAT1.2 in undifferentiated hESCs or mESCs (Figures 1D, 1E, S2A, and S2B).

We found previously that NEAT1 is strongly bound by TDP-43 in somatic cells (Tollervey et al., 2011), which led us to inquire
Figure 1. Alternative Polyadenylation of NEAT1 Induces Paraspeckle Formation upon Differentiation and Depletion of TDP-43 in Mouse and Human ESCs

(A and B) The number of paraspeckles analyzed by counting NEAT1_1, NEAT1_2 (A) or Neat1_1, Neat1_2 (B) double-labeled foci (based on single-molecule fluorescent in situ hybridization [smFISH] and criteria explained in Figure S1A).

(A) Undifferentiated hESCs, spontaneously differentiating cells, and BMP4-, CHIR99021-, WNT3A-, and retinoic acid (RA)-treated cells, promoting trophoblast, mesoderm, mesendoderm (primitive streak-like), and neuroectoderm fates, respectively.

(B) Undifferentiated mESCs and spontaneously differentiating iTDP-43-EGFP mESCs untreated or treated with doxycycline to ectopically express TDP-43; more than 250 (A) and more than 200 (B) cells analyzed per group, Mann-Whitney U test; **p < 0.001, ***p < 0.0001. Duration of treatment was as indicated.

(C) Venn diagram depicting differentially expressed genes in hESCs exposed to the indicated differentiation stimuli for 24 h relative to untreated cells (n = 2 biological replicates of global RNA-seq per condition; adjusted p < 0.01, Fisher’s exact test, fold change ≥ 4). Bottom: representative mapping of NEAT1 RNA-seq reads aligned to the two isoforms; global and nascent RNA-seq of undifferentiated hESCs and CHIR99021-treated hESCs (n = 2 biological replicates).

(D and E) Percentage and representative maximum projection photomicrographs of h/mESCs exhibiting NEAT1_1, NEAT1_2 (D) and Neat1_1, Neat1_2 (E) isoforms analyzed as above.

(F) Representative mapping of NEAT1 and Neat1 RNA-seq reads displaying samples from TDP-43 siRNA KD or control siRNA-treated hESCs (2 days) and untreated or tamoxifen-treated Tdp-43 KO mESCs (2 days, n = 3 biological replicates under pluripotency conditions).
parcially reduces its binding to polyadenylated RNAs upon para-
speckle formation.

TDP-43 Maintains Pluripotency and Enhances Somatic
Cell Reprogramming

The role of TDP-43 in suppressing paraspeckle formation in undifferentiated hESCs or mESCs and TDP-43 sequestration in differentiated cells suggested that it might be regulating pluripotency. To directly examine this, we extracted transcript abundance from 3' mRNA-seq of conditional Tdp-43 knockout (cTdp-43 KO) mESCs (Chiang et al., 2010; Figure S2C). This demonstrated that many general pluripotency genes, as defined by a previous study (Kalkan et al., 2017), were significantly downregulated, whereas genes that collectively classify trophoe-
derm differentiation were upregulated (Figure 3A; Table S3). The only significantly upregulated pluripotency gene was Lin28a, but this is consistent with the fact that expression of Lin28a increases immediately after the exit from naive pluripo-
tency (Kalkan et al., 2017). Destabilization of pluripotency was also evident by the loss of the dome-shaped compact colony morphology in cTdp-43 KO mESCs (Figure 3B). In contrast, over-
expression of TDP-43 promoted pluripotency, as indicated by prolonged expression of SSEA-1 and elevated NANOGEN upon spontaneous differentiation of mESCs (Figures 3C and 3D).

To analyze this further, we next tested whether overexpression of TDP-43 affects reprogramming of human secondary fibroblasts harboring a doxycycline-inducible OKSM (OCT4, KLF4, SOX2, C-MYC) gene cassette (Cacchiarelli et al., 2015; Figure 3E). We found that this results in a nearly 2-fold increase in the number of colonies expressing the pluripotency marker alkaline phosphatase (AP) and a 3-fold increase in cells expressing the pluripotency marker TRA-1-60 (Figures 3F–3I), indicating that TDP-43 facilitates reprogramming. Moreover, transient ectopic expression of TDP-43 during the initial phase of reprogramming of primary adult human fibroblasts led to a 5-fold increase in the number of induced pluripotent stem cell (iPSC) colonies (Figures 3J and 3K).

TDP-43 Regulates APA of Many Transcripts Important
for Pluripotency, Including SOX2

To further understand the mechanisms whereby TDP-43 pro-
motes pluripotency, we analyzed its global role in pA site selec-
tion, which is plausible given our previous findings regarding TDP-43's direct regulation of APA (Rot et al., 2017). We used 3' mRNA-seq to compare undifferentiated hESCs with meso-
dermal progenitors that were differentiated by activating the Wnt or β-catenin pathway using a GSK3 inhibitor (Blauwkamp et al., 2012). We confirmed the enrichment of mesoderm gene signature and downregulation of pluripotency genes in the pro-
genitor population by RNA-seq (Figure S3A; Table S4). We also confirmed the validity of the pA sites identified by 3' mRNA-
seq (Figures S3B–S3F). In genes that contained multiple pA sites, we defined the two pA sites with the most significant change between mesodermal progenitors and undifferentiated hESCs. Importantlay, we found a correlation between the pA site changes induced by differentiation or by TDP-43 knockdown (KD) in undifferentiated hESCs (Pearson’s correlation coefficient (r) = 0.62; Figure 4A; Table S5). This indicates that the decreased activity of TDP-43 upon early differentiation is a major cause of differentiation-induced APA.

To systematically identify the primary RNA motifs that contribute to the regulation of APA, we analyzed the enriched sequence motifs. We identified GU-rich motifs as the highest-
ranking sequences that are enriched around the pA sites that changed during hESC differentiation (Figure S4A). Notably, similar motifs were enriched around differential pA sites affected by KD of TDP-43 in undifferentiated hESCs and by cTdp-43 KO in mESCs (Figure S4A). Moreover, RNA maps of these motifs showed that they are located in the region upstream of the regulated pA sites (Figures S4B–S4D), where accessory regulatory sites typically reside (Di Giampartino et al., 2011). The similar pattern of motif enrichment for all conditions (Figure S4A) and the high correlation in processing changes upon differentiation or TDP-43 KD (Figure 4A) indicated that binding of TDP-43 close to the regulated pA sites could explain a large portion of the global changes in APA that take place upon differentiation of hESCs or mESCs.

To validate the direct binding of TDP-43 to the regulated pA sites, we conducted individual nucleotide-resolution UV cross-
linking and immunoprecipitation (iCLIP) experiments with TDP-
43 in hESCs (König et al., 2010). A direct role of TDP-43 in APA regulation was evident by the clear position-dependent dual enrichment of its binding sites around the pA sites that undergo
Figure 3. TDP-43 Maintains Pluripotency and Enhances Somatic Cell Reprogramming

(A) Volcano plot displaying gene expression fold changes and their respective statistical score (adjusted p value, Fisher’s exact test), comparing untreated and tamoxifen-treated cTdp-43 KO undifferentiated mESCs (n = 10 biological replicates per condition). Known pluripotency and trophectoderm markers are labeled red and blue, respectively (marked factors fulfilled p < 0.01).

(legend continued on next page)
APA following TDP-43 KD or KO, as evident by the RNA map showing the density of its crosslinking in hESCs (Figures S5A–S5E) or of GU-rich motifs (Figures S5F and S5G), consistent with the previously reported RNA map in HEK293 cells (Rot et al., 2017). Enriched binding of TDP-43 was detected close to repressed pA sites (within 100 nt) and further upstream and downstream of repressed sites (Figure S5F) in a pattern that was also conserved in mESCs (Figure S5G). Notably, TDP-43-dependent regulation of over 40% of the affected pA sites was conserved between hESCs and mESCs (Table S5). These analyses validated the high quality and positional precision of our data and identified the RNA targets that are directly regulated by TDP-43 according to position-dependent principles.

Notably, transcripts exhibiting APA upon cTdp-43 KO were significantly enriched for the GO-term “stem cell maintenance” (Figure 4B; Table S5). Moreover, TDP-43 dictated the APA significantly enriched for the GO-term “stem cell maintenance” by TDP-43 according to position-dependent principles.

To decipher whether TDP-43 directly regulates NEAT1 processing, we analyzed the RNA-binding profile of TDP-43 by icerIP (Figures S5A–S5G). This showed that TDP-43 binds to evolutionarily conserved clusters of GU-rich motifs upstream of the pA signal in mESCs (Figure S5D). This pattern agrees with the enhancing effect of TDP-43 when binding more than 100 bp upstream of the pA site in Sox2 (Figure 4E). To validate the direct effect, we used CRISPR-Cas9 editing of mESCs to excise a 100-nt GU-rich binding site of TDP-43 that is located 105 nt upstream of the internal Neat1 pA site (Figure 5E; Neat1_1100nt mESCs). Importantly, undifferentiated Neat1_1100nt mESCs exhibited high expression of the Neat1_2 isoform (Figure 5F) and paraspeckles (Figure 5G), which were induced to a comparable amount as in cTdp-43 KO or Neat1_3pA mESCs. Moreover, we removed growth factors promoting pluripotency from Neat1_1100nt and control mESC cultures to induce paraspeckles while also ectopically expressing iTDP-43-EGFP. The capacity of TDP-43-EGFP to repress the expression of Neat1_2 under these conditions was reduced 2-fold in Neat1_1100nt cells (Figure S5J). Finally, Neat1_1100nt mESCs exhibited an increased differentiation propensity, as evident by the decline of NANOG during spontaneous differentiation (Figure S5K). We concluded that TDP-43 binding upstream of the Neat1 pA site enhances the production of a dramatic induction of Sox2 in primitive streak progenitors (Figure 4L). These results show that TDP-43 promotes pluripotency by APA-mediated regulation of pluripotency factors, including Sox2.

TDP-43 Represses Formation of Paraspeckles by Regulating APA of NEAT1

Beyond direct regulation of pluripotency genes via APA, we sought to understand how TDP-43 regulates paraspeckle formation via APA of NEAT1. Given that the NEAT1 isoform switch takes place upon differentiation and KD of TDP-43 (Figure 1), we speculated that downregulation of TDP-43 coincides with and, therefore, contributes to the isoform switch during differentiation. Indeed, decreased expression of TDP-43 was evident in a published RNA-seq dataset of hESCs differentiated into developmental progenitors (Gifford et al., 2013), which was inverse to the upregulation of NEAT1_2 (Figure 5A). Moreover, we observed significantly lower levels of TDP-43 transcript and protein upon 3 days of differentiation (Figures 5B and 5C; Figure S5H). Similarly, the expression of NEAT1 was inversely correlated with pluripotency genes and TDP-43 during the reprogramming of human fibroblasts to iPSCs (Figure S5I).

To understand how TDP-43 promotes Sox2 expression, we created a minigene reporter containing the 3’ UTR of Sox2 located downstream of EGFP (Figure 4G). Flow cytometry showed that overexpression of TDP-43 in HEK293T cells increased the GFP signal in the presence of miR-21, which could target the long isoform of Sox2 (Figure 4H). Mutating the proximal pA site in the Sox2 3’ UTR made it insensitive to TDP-43, confirming that TDP-43 regulates Sox2 expression through APA (Figure 4I). Finally, mutating the miR-21 binding site in the reporter construct, which resides in the long isoform, ablated the effect of TDP-43 on EGFP expression (Figure 4J). Importantly, upregulation of miR-21 took place upon primitive streak differentiation (Figure 4K), which agrees with the finding that deletion of its binding site in Sox2 by genome editing led to a dramatic induction of Sox2 in primitive streak progenitors (Figure 4L). These results show that TDP-43 promotes pluripotency by APA-mediated regulation of pluripotency factors, including Sox2.
Figure 4. TDP-43 Regulates APA of Genes Important for Pluripotency, Including Sox2

(A) Scatterplot displaying relative changes of pA sites upon differentiation of hESCs (Figures S3A and S3B) and changes accruing by knocking down TDP-43 in undifferentiated hESCs for 48 h (n = 4 control short hairpin [shCTRL] and n = 8 shTDP-43 using 2 TDP-43-targeting short hairpin RNAs [shRNAs] with 4 replicates for each; adjusted p < 0.05, Fisher’s exact test). Linear regression (gray line) and the 90% confidence interval region (light blue) are shown (Pearson’s correlation (legend continued on next page))
polyadenylated Neat1_2, which decreases the production of Neat1_2 and assembly of paraspeckles upon differentiation.

**Neat1_2 Is Required for Efficient Early Differentiation by Cross-Regulation with TDP-43**

To analyze whether Neat1 functionally regulates the differentiation of ESCs, we excised the triple helix (TH) region at the 3' end of the gene, which is required for stabilizing Neat1_2 in somatic cells (Wilusz et al., 2012; Yamazaki et al., 2018). Homozygous TH deletions in mESCs significantly reduced the number of paraspeckles in differentiated cells (Figure 6A; Figure S6A). Notably, the in vitro pluripotency phenotype of Neat1_1TH mESCs was significantly prolonged during differentiation, as evident by higher expression of the pluripotency markers SSEA-1 and NANOG (Figures 6B and 6C), mimicking the phenotype of TDP-43 overexpression (Figures 3C and 3D).

To further assess the function of Neat1_2 during embryonic development, we used tetraploid (4n) mouse embryo ↔ 2n mESC aggregations for generating chimera embryos exclusively derived from mESCs, whereas extra-embryonic lineages were derived from the 4n cells (Figure 6D; Eakin and Hadjantonakis, 2006). We aggregated 4n membrane-targeted, tdTomato-expressing, 2- to 4-cell-stage embryos with wild-type (WT), Neat1_2PA, or Neat1_1TH mESCs, which we generated by editing IDG3.2 mESCs, a line that is routinely used for generating full-term chimeras (Engert et al., 2013). Embryonic day 7.75 (E7.75) to E8.0 embryos derived from Neat1_1TH mESCs were abnormal and exhibited defective patterning of axial mesoderm and the node, as indicated by scattered expression of Sox2 and Brachyury (Figures 6E and 6F; Figures S6B and S6C). Scattered expression of Sox2 in Figure S6E also suggested that the identity of definitive and visceral endoderm was perturbed (Figure S6C).

In contrast, we did not observe morphological abnormalities of chimera embryos derived from Neat1_2PA or WT mESCs (Figures 6E and 6F; Figures S6D and S6E), which efficiently form paraspeckles (Figure 1D). We also looked for phenotypic effects of inducing Neat1_2 in undifferentiated, pluripotent NEAT1_2pA hESCs. Comparison with the parental cells showed that a cohort of mesoderm and endoderm markers were already expressed in the undifferentiated state, plausibly in subpopulations, despite the cells being maintained in pluripotency conditions (Figure S6F). Taken together, these results demonstrate that paraspeckle formation facilitates the differentiation of pluripotent cells and embryonic patterning.

Last, we analyzed the interdependence of TDP-43 and Neat1_2 in the regulation of early spontaneous differentiation. We manipulated the two factors concomitantly by deleting the endogenous TH region in Neat1_2 in the cTdp-43 KO mESC line. Paraspeckles are normally increased upon inducible KO of Tdp-43 by tamoxifen treatment of cTdp-43 KO mESCs (Figure 1E), whereas Neat1_1TH mESCs are defective in the induction of paraspeckles (Figure 6A; Figure S6A). We assessed whether deficient induction of paraspeckles upon tamoxifen treatment of the cTdp-43 KO Neat1_1TH mESCs affects any of the cellular responses to TDP-43 depletion. Indeed, the key pluripotency proteins SOX2 and NANOG were downregulated to a lesser extent upon spontaneous differentiation and tamoxifen treatment of cTdp-43 KO Neat1_1TH mESCs with deficient paraspeckle induction compared with cTdp-43KO mESCs Figures 6G–6J). Global transcriptome analysis revealed the same trend, evident by maintained expression of pluripotency genes and reduced expression of early differentiation genes, including Cdx2 and Brachyury (Figure 6K; Table S6). This was supported by the higher proliferative capacity of tamoxifen-treated differentiating cTdp-43 KO Neat1_1TH mESCs compared with Neat1 WT mESCs (Figure 6L). Strikingly, we observed a dramatic reduction in the number of APA-regulated genes in tamoxifen-treated cTdp-43KO Neat1_1TH mESCs as compared with control cells where induction of cTdp-43KO induces paraspeckles (Figure S6M; Table S6). Thus, decreased abundance of TDP-43 promotes dissolution of pluripotency largely through APA regulation, but induction of Neat1_2 is required for efficient APA regulation. Collectively, we conclude that cross-regulation between Neat1 and TDP-43 is essential for their efficient reciprocal roles in the pluripotency-differentiation transition.
Network modules that promote efficient cell state transitions are inherent to differentiation, regulation of organismal development, homeostasis, and regeneration (Alon, 2007). Cross-regulation between regulatory factors can promote lineage bifurcation and has been associated with transcription factors, microRNAs, and signaling cascades.

**DISCUSSION**

Network modules that promote efficient cell state transitions are inherent to differentiation, regulation of organismal development, homeostasis, and regeneration (Alon, 2007). Cross-regulation between regulatory factors can promote lineage bifurcation and has been associated with transcription factors, microRNAs, and signaling cascades.
Changes in APA have been observed previously during differentiation processes that include myogenesis, adipogenesis, and embryogenesis (Di Giammartino et al., 2011; Hoque et al., 2013). Moreover, the core elements of the cleavage and polyadenylation complex, FIP1 (Lackford et al., 2014) and NUDT21 (Brumbaugh et al., 2016) have been implicated in the regulation of APA during ESC differentiation. Here we used iCLIP and motif analysis to show how TDP-43 regulates APA of specific transcripts through binding to RNA motifs around the regulated pA sites (Figures S4 and S5). Therefore, TDP-43 has a conserved role in promoting the production of short Sox2 transcript, which is required for efficient expression of this pluripotency and reprogramming factor. It also promotes the short isoform of Neat1_1, which suppresses paraspeckle formation. This role of TDP-43 in early cell fate transitions is in line with the past observation that it is essential for murine development (Sephton et al., 2010).

When the pendulum swings in the direction of differentiation, the abundance of TDP-43 decreases to a level permissive for an isoform switch in NEAT1, resulting in the production of NEAT1_2 that scaffolds paraspeckles. It has been previously shown that efficient interaction of TDP-43 and several other RBPs with NEAT1_2 occurs as a result of phase separation into paraspeckles (Maharana et al., 2018; Yamazaki et al., 2018). Here we have shown that induction of NEAT1_2 in ESCs partly sequesters TDP-43 and other paraspeckle RBPs away from mRNAs, most likely because of their recruitment into the paraspeckle compartment. This sequestration could contribute to the decreased activity of TDP-43 upon early differentiation, facilitating the global shift in APA and, therefore, exit from pluripotency, which could partly explain how Neat1_2 promotes early differentiation. It is likely that Neat1 also acts through additional mechanisms, such as the recently reported role of NEAT1_2 in scaffolding microprocessor and other RBPs into paraspeckles to regulate microRNA biogenesis (Jiang et al., 2017), and Neat1_1 on its own could form “microspeckles,” whose functions are currently unknown (Li et al., 2017). Thus, our study indicates that formation of a membraneless compartment can be regulated through the processing of its RNA scaffold to coordinate a specific developmental transition.

We have characterized developmental phenotypes of embryonic cells lacking the TH of Neat1, which have a reduced number of paraspeckles because of destabilized Neat1_2. We used the tetraploid complementation approach, which normally generates viable embryos from ESCs and is used to investigate the role of factors of interest in embryonic tissues (Eakin and Hadjantonakis, 2006). This method enables classifications of embryonic versus extraembryonic phenotypes by circumventing pre-blastocyst developmental arrest, which has been observed upon depletion of Neat1_2 in a recent study (Hupaloowska et al., 2018). Tetraploid complemented embryos derived from Neat1_1TH mESCs have inefficient patterning during gastrulation. Notably, we observed elevated expression of the transcription factor Cdx2 as the major change that distinguishes the response of Neat1_1TH and WT mESCs to loss of TDP-43 (Figure 6K), which agrees with the previously observed role of Neat1 in regulating Cdx2 (Hupaloowska et al., 2018). The findings of both studies are also consistent with previous studies that observed decreased litter size, aberrant mammary gland morphogenesis, and smaller pups (Nakagawa et al., 2014; Standaert et al., 2014) in Neat1 KO mice, which lack paraspeckles (Nakagawa et al., 2011). This phenotype of Neat1 KO mice is surprisingly mild, given the apparent phenotypes observed in the study, which indicates that the function of Neat1 might be compensated by additional factors that assist cell fate transitions during in vivo murine development.

Beyond the mechanisms of early development, our findings are likely also relevant also for other types of cell fate transitions because both TDP-43 and NEAT1 are broadly expressed (Shelkovnikova et al., 2018). Moreover, amyotrophic lateral sclerosis (ALS)-causing mutations in TDP-43 affect its phase separation.
properties (Conicella et al., 2016), which could affect its localization into paraspeckles (Yamazaki et al., 2018). Moreover, increased formation of paraspeckles has been observed in motor neurons of patients with ALS (Nishimoto et al., 2013). Therefore, cross-regulation of TDP-43 and NEAT1 could play a role not just in development but also in diseases such as ALS and cancer, where abnormal RNA granules involving TDP-43 and NEAT1 are a common feature (Adriaens et al., 2016; Hennig et al., 2015; Nishimoto et al., 2013; Tollervey et al., 2011).

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**
  - Cell culture
  - Generation of h/mESC lines harboring edited NEAT1/Neat1
  - Generation of HAP-1 cells harboring edited NEAT1
  - Overexpression of TDP-43
  - Knockdown of TDP-43
  - Plasmids
  - Reprogramming of human fibroblasts
- **ANIMALS**
- **METHOD DETAILS**
  - RNA preparation and qRT-PCR assays
  - RNA sequencing
  - Western blotting
  - Immunofluorescence
  - Flow cytometry
  - Subcellular fractionation of hESCs
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Single molecule (sm)FISH and paraspeckle quantification
  - RNA-FISH and immunofluorescence
  - iCLIP protocol and its analysis
  - RNA-Seq differential expression
  - Identification of polyA sites
  - Analysis of alternative polyadenylation (APA)
  - Visualizing position-dependent polyA site regulation using RNA-maps
  - mRNA/RBP occupancy (mRNA interactome capture)
  - FASP digest
  - Mass spectrometry
  - Analysis of mRNA- interactome
- **DATA AND SOFTWARE AVAILABILITY**

Supplemental Information

Supplemental Information can be found online at https://doi.org/10.1016/j.molcel.2019.03.041.

Acknowledgments

The authors would like to thank Alfredo Castello and Rickie Patani for critical reading of the manuscript. For generous reagent gifts, we are grateful to Derk ten Berge, Philip C. Wong, Shinichi Nakagawa, Michael Kyba, Dieter Edbauer, and Bettina Schmidt. We would like to thank to Anna Pertek and Tajda Klobucar for technical assistance. We are grateful to Nejc Haberman for sequence motif identification, Igor Ruiz de los Mozos for iCount analysis, and Michael Ziller and Joel Ryan for methodological advice. This work was supported by the European Research Council (206726-CLIP and 617837-Translate to J.U.), Slovenian Research Agency (P4-0127, JS-6789, J3-8201, and J3-9263 to B.R.), an IMPRS Max Planck Society fellowship (to M.M.), the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001002), the UK Medical Research Council (FC001002), and the Wellcome Trust (FC001002), and Seed Funding Award by the UHU Network (UWA, Helmholz, UCL, to M.D.).

Author Contributions

M.M. conceived the study and performed majority of the experiments apart from smFISH, which was carried out and analyzed by M.G. M.G. also generated modified hESC lines. M.G., E.R., and M.M. performed flow cytometry of genetically edited m/hESC lines. S.S. performed chimera experiments under the supervision of H.L. T.Y. generated ΔUG and UG60 HAP1 lines and performed imaging experiments with them under the supervision of T.H., E.R., T.L., D.C., and A.M. contributed to the reprogramming experiments. F.C.Y.L. and M.M. performed mRNA-RBP occupancy experiments, assisted by M.P., and S.M.H. and J.M.-P. oversaw the MS analysis. D.S. performed nascent RNA-seq. G.R. and M.M. analyzed the data, with contributions from C.v.M. and B.R. M.D. and J.U. jointly supervised the study. M.M., J.U., and M.D. wrote the manuscript.

Declaration of Interests

The authors declare no competing interests.

Received: April 29, 2018
Revised: February 12, 2019
Accepted: March 28, 2019
Published: April 29, 2019

References

Adriaens, C., Standaert, L., Barra, J., Latil, M., Verfaille, A., Kaley, P., Boexck, B., Wijnhoven, P.W.G., Radaelli, E., Vermi, W., et al. (2016). p53 induces formation of NEAT1 IncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity. Nat. Med. 22, 861–866.

Akhtar, M.N., Bukhari, S.A., Fazal, Z., Qamar, R., and Shahmuradov, I.A. (2010). POLYAR, a new computer program for prediction of poly(A) sites in human sequences. BMC Genomics 11, 646.

Alon, U. (2007). An Introduction to Systems Biology: Design Principles of Biological Circuits, Third Edition (Chapman and Hall/CRC).

Baltz, A.G., Munschauer, M., Schwahnhauser, B., Vasilie, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., et al. (2012). The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol. Cell 46, 674–690.

Blake, J.A., Christie, K.R., Dolan, M.E., Drabkin, H.J., Hill, D.P., Ni, L., Sitrinov, D., Burgess, S., Buza, T., Gresham, C., et al.; Gene Ontology Consortium (2015). Gene Ontology Consortium: going forward. Nucleic Acids Res. 43, D1049–D1056.

Blauwkamp, T.A., Ngiam, S., Ardehali, R., Weissman, I.L., and Nusse, R. (2012). Endogenous Wnt signaling in human embryonic stem cells generates an equilibrium of distinct lineage-specified progenitors. Nat. Commun. 3, 1070.

Brumbaugh, J., Di Stefano, B., Wang, X., Borkent, M., Forouzmand, E., Clowers, K.J., Ji, F., Schwarz, B.A., Kalocsay, M., Elledge, S.J., et al. (2018). Nud21 Controls Cell Fate by Connecting Alternative Polyadenylation to Chromatin Signaling. Cell 722, 106–120.e21.

Budini, M., Romano, V., Quadri, Z., Buratti, E., and Baralle, F.E. (2015). TDP-43 loss of cellular function through aggregation requires additional structural
determinants beyond its C-terminal Q/N prion-like domain. Hum. Mol. Genet. 24, 9–20.

Cacciarielli, D., Trapnell, C., Ziller, M.J., Soumillon, M., Cesana, M., Karnik, R., Donaghey, J., Smith, Z.D., Ratansiriratwoot, S., Zhang, X., et al. (2015). Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of Induced Pluripotency. Cell 162, 412–424.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strie, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., et al. (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteomes. Cell 749, 1393–1406.

Castello, A., Horos, R., Strein, C., Fischer, B., Eichelbaum, K., Steinmetz, L.M., Krijgsveld, J., and Hentze, M.W. (2013). System-wide identification of RNA-binding proteins by interactome capture. Nat. Protoc. 8, 491–500.

Chen, L.-L., and Carmichael, G.G. (2009). Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. Mol. Cell 35, 467–478.

Chiang, P.M., Ling, J., Jeong, Y.H., Price, D.L., Aja, S.M., and Wong, P.C. (2010). Deletion of TDP-43 down-regulates Tbc1d1, a gene linked to obesity, and alters body fat metabolism. Proc. Natl. Acad. Sci. USA 107, 16320–16324.

Concilla, A.E., Mittal, J., Fawzi, N.L., Concilla, A.E., Zerhe, Z., Mittal, J., and Fawzi, N.L. (2016). ALS Mutations Disrupt Phase Separation Mediated by α-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. Structure 24, 1537–1549.

Derti, A., Garrett-Engele, P., McIsaac, K.D., Stevens, R.C., Siriam, S., Chen, R., Rohl, C.A., Johnson, J.M., and Babak, T. (2012). A quantitative atlas of polyadenylation in five mammals. Genome Res. 22, 1173–1183.

Di Giannantonio, D.C., Nishida, K., and Manley, J.L. (2011). Mechanisms and consequences of alternative polyadenylation. Mol. Cell 43, 853–866.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.

Eakin, G.S., and Di Giannantonio, A.K. (2006). Production of chimeras by aggregation of embryonic stem cells with diploid or tetraploid mouse embryos. Nat. Protoc. 1, 1145–1153.

Engert, S., Burtscher, I., Liao, W.P., Dulev, S., Schotta, G., and Lickert, H. (2013). Wnt/Catenin Signaling regulates Sox17 expression and is essential for organizer and endoderm formation in the mouse. Development 140, 3128–3138.

Enver, T., Pera, M., Peterson, C., and Andrews, P.W. (2009). Stem cell states, fates, and the rules of attraction. Cell Stem Cell 5, 467–478.

Fawzi, N.L. (2016). ALS Mutations Disrupt Phase Separation Mediated by α-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. Structure 24, 1537–1549.

Huppertz, I., Attig, J., D’Ambrogio, A., Easton, L.E., Sibley, C.R., Sugimoto, Y., Tajnik, M., König, J., and Ule, J. (2014). ICLIP: protein-RNA interactions at nucleotide resolution. Methods 65, 274–287.

Huang, S., Shao, C., Wu, Q.J., Chen, G., Zhou, J., Yang, B., Li, H., Gou, L.T., Zhang, Y., Wang, Y., et al. (2017). NEAT1 scaffolds RNA-binging proteins and the Microprocessor to globally enhance pri-miRNA processing. Nat. Struct. Mol. Biol. 24, 816–824.

Kalkan, T., Olova, N., Roode, M., Mulis, 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.

Kawaguchi, T., Tanigawa, A., Naganuma, T., Otkaawa, Y., Souquere, S., Pierron, G., and Hirose, T. (2015). SWI/SNF chromatin-remodeling complexes function in noncoding RNA-dependent assembly of nuclear bodies. Proc. Natl. Acad. Sci. USA 112, 4300–4309.

König, J., Zamack, K., Rot, G., Curk, T., Kaykci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). ICLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat. Struct. Mol. Biol. 17, 909–915.

Kyba, M., Perlingeiro, R.C.R., and Daley, G.Q. (2002). HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. Cell 109, 29–37.

Lackford, B., Yao, C., Charles, G.M., Weng, L., Zheng, X., Choi, E.A., Xie, X., War, J., Xing, Y., Freudenberg, J.M., et al. (2014). Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. EMBO J. 33, 878–889.

Li, R., Harvey, A.R., Hodgetts, S.J., and Fox, A.H. (2017). Functional dissection of NEAT1 using genome editing reveals substantial localization of the NEAT1_1 isoform outside paraspeckles. RNA 23, 872–881.

Ling, J.P., Pletnikova, O., Troncoso, J.C., and Wong, P.C. (2015). TDP43 repression of nonconserved cryptic exons is compromised in ALS-FTD. Science 349, 650–655.

Maharana, S., Wang, J., Papadopoulos, D.K., Richter, D., Pozniakovsky, A., Poser, I., Bickle, M., Rizk, S., Guillen-Boixet, J., Franzmann, T.M., et al. (2018). RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science 360, 918–921.

Mao, Y.S., Sunwoo, H., Zhang, B., and Spector, D.L. (2011). Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. Nat. Cell Biol. 13, 95–101.

Merj, L., Ueffing, M., Hauck, S.M., and von Toerne, C. (2012). Direct comparison of MS-based label-free and SILAC quantitative proteome profiling strategies in primary retinal Muller cells. Proteomics 12, 1902–1911.

Muzumdar, M.D., Tasic, B., Miyamichi, K., Li, L., and Luo, L. (2007). A global double-fluorescent Cre reporter mouse. Genetics 173, 593–605.

Nagamura, T., Nakagawa, S., Tanigawa, A., Sasaki, Y.F., Goshima, N., and Hirose, T. (2012). Alternative 3′-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. EMBO J. 31, 4020–4034.

Nakagawa, S., Naganuma, T., Shioi, G., and Hirose, T. (2011). Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. J. Cell Biol. 193, 31–39.

Stickel, S., Shimada, M., Yanaka, T., Itagaki, M., Arai, T., Takahashi, E., Fujita, Y., Fujimori, T., Standaert, D., Marine, J.-C., and Hirose, T. (2014). The IncRNA Neat1 is required for corpus luteum formation and the establishment of pregnancy in a subpopulation of mice. Development 141, 4618–4627.

Nishimoto, Y., Nakagawa, S., Hirose, T., Okano, H.J., Takao, M., Shibata, S., Suyama, S., Kusakari, K., Imai, T., Murayama, S., et al. (2013). The long non-coding RNA nuclear-enriched abundant transcript 1.2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis. Mol. Brain 6, 31.

Hupalowska, A., Jedrusik, A., Zhu, M., Bedford, M.T., Glover, D.M., and Zernicka-Goetz, M. (2018). CARMA1 and paraspeckles regulate pre-implantation mouse embryo development. Cell 175, 1902–1916.e13.

Huppertz, I., Attig, J., D’Ambrogio, A., Easton, L.E., Sibley, C.R., Sugimoto, Y., Tajnik, M., König, J., and Ule, J. (2014). ICLIP: protein-RNA interactions at nucleotide resolution. Methods 65, 274–287.
Prasanth, K.V., Prasanth, S.G., Xuan, Z., Hearn, S., Freier, S.M., Bennett, C.F., Zhang, M.O., and Spector, D.L. (2005). Regulating gene expression through RNA nuclear retention. Cell 123, 249–263.

Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. Nat. Methods 5, 877–879.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140.

Rot, G., Wang, Z., Huppertz, I., Modic, M., Lence, T., Hallegger, M., Haberman, N., Curk, T., von Mering, C., and Ule, J. (2017). High-Resolution RNA Maps Suggest Common Principles of Splicing and Polyadenylation Regulation by TDP-43. Cell Rep. 19, 1056–1067.

Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E.E., Nitsch, R., and Wulczyn, F.G. (2008). A feedback loop comprising lin-28 and let-7 maturation during neural stem-cell commitment. Nat. Cell Biol. 10, 987–993.

Schwenk, B.M., Hartmann, H., Serdaroglu, A., Schludi, M.H., Hornburg, D., Meissner, F., Orozco, D., Colombo, A., Tähnroth, S., Michaelisen, M., et al. (2016). TDP-43 loss of function inhibits endosomal trafficking and alters tropic signaling in neurons. EMBO J. 35, 2350–2370.

Sephton, C.F., Good, S.K., Atkin, S., Dewey, C.M., Mayer, P., 3rd, Herz, J., and Yu, G. (2010). TDP-43 is a developmentally regulated protein essential for early embryonic development. J. Biol. Chem. 285, 8826–8834.

Shelkovnikova, T.A., Kulharsky, M.S., An, H., Dimasi, P., Alexeeva, S., Shabir, O., Heath, P.R., and Buchman, V.L. (2018). Protective paraspeckle hyper-assembly downstream of TDP-43 loss of function in amyotrophic lateral sclerosis. Mol. Neurodegener 13, 30.

Shervtsov, S.P., and Dnundr, M. (2011). Nucleation of nuclear bodies by RNA. Nat. Cell Biol. 13, 167–173.

Standeart, L., Adriaens, C., Radaelli, E., Van Keymeulen, A., Blanpain, C., Hirose, T., Nakagawa, S., and Marine, J.C. (2014). The long noncoding RNA Neat1 is required for mammary gland development and lactation. RNA 20, 1844–1849.

Tollervey, J.R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., König, J., Hortobagyi, T., Nishimura, A.L., Zupunski, V., et al. (2011). Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. Nat. Neurosci. 14, 452–458.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105–1111.

Trcek, T., Lionnet, T., Shroff, H., and Lehmann, R. (2017). mRNA quantification using single-molecule FISH in Drosophila embryos. Nat. Protoc. 12, 1326–1348.

Tsankov, A.M., Akopian, V., Pop, R., Chetty, S., Gifford, C.A., Daheron, L., Tsankova, N.M., and Meissner, A. (2015). A qPCR ScoreCard quantifies the differentiation potential of human pluripotent stem cells. Nat. Biotechnol. 33, 1182–1192.

Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., Gaasterland, T., Blencowe, B.J., and Darnell, R.B. (2006). An RNA map predicting Nova-dependent splicing regulation. Nature 444, 580–586.

West, J.A., Mito, M., Kurosaka, S., Takumi, T., Tanegashima, C., Chuo, T., Yanaka, K., Kingston, R.E., Hirose, T., Bond, C., et al. (2016). Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization. J. Cell Biol. 214, 817–830.

Wilusz, J.E., JinBaptiste, C.K., Lu, L.Y., Kuhn, C.D., Joshua-Tor, L., and Sharp, P.A. (2012). A triple helix stabilizes the 3’ ends of long noncoding RNAs that lack poly(A) tails. Genes Dev. 26, 2392–2407.

Wisniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. Nat. Methods 6, 359–362.

Witten, J.T., and Ule, J. (2011). Understanding splicing regulation through RNA splicing maps. Trends Genet. 27, 89–97.

Yamazaki, T., Souquere, S., Chuo, T., Kobelke, S., Chong, Y.S., Fox, A.H., Bond, C.S., Nakagawa, S., Pierron, G., and Hirose, T. (2018). Functional Domains of NEAT1 Architectural IncRNA Induce Paraspeckle Assembly through Phase Separation. Mol. Cell 70, 1038–1053.e7.
**STAR★METHODS**

**KEY RESOURCES TABLE**

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Antibodies          |        |            |
| Sox2 monoclonal antibody (D6D9) Alexa 647-conjugated | Cell Signaling | Cat no. 5067; RRID: AB_10828229 |
| Sox2 monoclonal Antibody | Cell Signaling | Cat no. 2748S; RRID: 823640 |
| Nanog monoclonal Antibody | Cell Signaling | Cat no. 8822; RRID: AB_11217637 |
| TDP-43 polyclonal Antibody | Sigma-Aldrich | Cat no. SAB4200006; RRID: AB_10610848 |
| TDP-43 polyclonal Antibody | Proteintech Group | Cat no. 10782-2-AP; RRID: AB_615042 |
| GFP polyclonal Antibody | Molecular Probes/Thermo Fisher | Cat no. A6455; RRID: AB_221570 |
| SFPO monoclonal Antibody | Thermo Fisher Scientific | Cat no. MA1-25325; RRID: AB_2186931 |
| NONO polyclonal Antibody | Abcam | Cat no. ab70335; RRID: AB_1269576 |
| CPEB1 polyclonal Antibody | Cell Signaling | Cat no. 13583 |
| PSPC1 rabbit polyclonal antibody | Naganuma et al., 2012 |
| SRIF3 polyclonal antibody | ProteinTech Group | Cat no. 10916-1-AP; RRID: AB_2185064 |
| TNPO1 monoclonal antibody | Abcam | Cat no. ab10303; RRID: AB_2206878 |
| U2AF2 monoclonal antibody | Sigma-Aldrich | Cat no. U4758; RRID: AB_262122 |
| Brachyury N-19 goat polyclonal antibody | Santa Cruz Biotechnologies | Cat no. sc17743; RRID: AB_634980 |
| Foxa2 (HNF-3b) M-20 goat polyclonal | Santa Cruz Biotechnologies | Cat no. sc-6554; RRID: AB_2262810 |
| RFP rabbit polyclonal | Rockland | Cat no. 600-401-379S; RRID: AB_2209751 |
| Donkey anti-rabbit IgG 555 | Invitrogen | Cat no. A31572; RRID: AB_162543 |
| Donkey anti-goat IgG 488 | Invitrogen | Cat no. A11055; RRID: AB_2534102 |
| Histone H3 polyclonal Antibody | Abcam | Cat no. ab1791; RRID: AB_302613 |
| Beta Actin polyclonal Antibody | New England Biolabs/Cell Signaling antibody | Cat no. 8H10D10; RRID: AB_2242334 |
| RNAPolII-S5,S2 | Abcam | Cat no. ab103968; RRID: AB_2687918 |
| DyLight 650-conjugated anti SSEA-1 (clone MC-480) | Thermo Fisher | Cat no. MA1-022-D650; RRID: AB_2536696 |
| Alexa Fluor 647-conjugated anti TRA-1-60 | BD Bioscience | Cat no. 560122; RRID: AB_1645448 |
| Alexa Fluor 647 Dk anti-Gt IgG (H+L) | Thermo Fisher | Cat no. A11058; RRID: AB_2534105 |
| Goat anti-rabbit IgM-HRP | Santa Cruz Biotechnology | Cat no. sc-2030; RRID: AB_631747 |
| Goat anti-mouse IgG HRP-conjugated | Dianova | Cat no. 115-035-003; RRID: AB_10015289 |

**Chemicals, Peptides, and Recombinant Proteins**

| (Z)-4-Hydroxytamoxifen | Sigma-Aldrich | H7904 |
| Neurobasal Medium | Thermo Fisher | 21103049 |
| DMEM F12 | Thermo Fisher | 11320033 |
| N2 | Thermo Fisher | 17502048 |
| B27 | Thermo Fisher | 17504044 |
| Glutamax | Thermo Fisher | 35050 |
| 1% nonessential amino acids | Thermo Fisher | 1140050 |
| 2-mercaptoethanol | Thermo Fisher | 31350-010 |
| MEK inhibitor PD0325901 | Axon Medchem | 1408 |
| GSK3 inhibitor CHIR99021 | Tocris | 4953/50 |
| LIF | produced in house | SV30160.03HI |
| Trypsin-EDTA | Thermo Fisher | 25200056 |
| HyClone Fetal Bovine Serum (FBS) | GE Healthcare | SV30160.03HI |
| mTeSR1 medium | Stem Cell Technologies | 058-50 |

(Continued on next page)
**Continued**

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Matrigel            | Corning| 356234     |
| Gentle Cell Dissociation Reagent | StemCell Technologies | 7174 |
| KnockOut Serum Replacement | Thermo Fisher | 10828028 |
| DMEM high glucose   | Thermo Fisher | 11965092 |
| Blue Sepharose 6 Fast Flow/ HisTrap HP columns | GE LifeSciences | 17-0412-01 |
| recombinant WNT3A   | D. Ten Berge | N/A       |
| recombinant BMP4    | R&D systems | 314-BP |
| Accutase            | Sigma-Aldrich | A6984 |
| Y-27632 ROCK inhibitor | R&D systems | 1254/10 |
| RPMI1640            | Thermo Fisher | 11875093 |
| L-Glutamine         | Thermo Fisher | 25030081 |
| B27 supplement without insulin | Thermo Fisher | A1895601 |
| Lipofectamine 3000  | Thermo Fisher | L3000008 |
| Lipofectamine RNAiMAX | Thermo Fisher | 13778150 |
| Puromycin           | Sigma-Aldrich | P8833-10MG, |
| blasticidin         | Invivogen | ant-bl-05 |
| G418                | Invivogen | ant-gn-01 |
| Trizol LS           | Life Technologies | 10296028 |
| QIAzol              | QIAGEN | 79306     |
| Zymo Direct-zol     | Zymogen | R2052     |
| RNeasy Mini Kit     | QIAGEN | 74104     |
| SuperScript III Reverse Transcriptase | Thermo Fisher | 18080085 |
| RNase H, recombinant | New England Biolabs | M0297S |
| miRNeasy Micro Kit  | QIAGEN | 217084    |
| KOD FX Neo enzyme   | Toyobo  | KFX-201    |
| Taqman assay master mixl | Thermo Fisher | 4369514 |
| Q5® High-Fidelity 2X Master Mix | Thermo Fisher | M0492 |
| miScript PCR Starter Kit | QIAGEN | 218193 |
| formamide           | Calbiochem | 4610-OP  |
| Dextran Sulfate     | VWR     | 9011-18-1  |
| competitor tRNA from E.coli | Roche Diagnostics | 10109541001 |
| methanol-free formaldehyde | Thermo Fisher | 28906 |
| digoxigenin (DIG) RNA labeling kit | Roche | 11175025910 |
| 2-well µ-Slides     | Ibidi   | 80286     |
| 8-chamb Slides      | Ibidi   | 80826     |
| DAPI                | New England Biolabs | 9071S |
| SYTOX Blue Dead Cell Stain, for flow cytometry | Thermo Fisher | S34857 |
| UltraPure BSA       | Thermo Fisher | AM2616 |
| vanadyl–ribonucleoside complex | New England Biolabs | S1402S |
| Insulin (human recombinant - yeast) | Life Tech | 12585-014 |
| Transferrin (human recombinant) | Sigma-Aldrich | T3705-1G |
| Sodium selenite     | Sigma-Aldrich | S5261-10G |
| Ascorbic acid 2-phosphate | Sigma-Aldrich | A8960-5G |
| FGF2 (human recombinant - E. coli) | Peprotech | 100-18B (1 mg) |
| TGFβ (human recombinant - E. coli) | Peprotech | 100-21 (100 ug) |
| TURBO Dnase         | Thermo Fisher | am2238 |
| RNeasy MinElute RNA cleanup | QIAGEN | 74204 |

(Continued on next page)
**Continued**

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| protease inhibitors | Merck Millipore | 539134 |
| phosphatase inhibitors | Sigma-Aldrich | 4906837001 |
| proteasome inhibitor | Sigma -Aldrich | MG132 |
| Mini-PROTEAN TGX Stain Free Gels, 4-15% | Bio-Rad Laboratories | 456-8086 |
| Clarity Western ECL Substrate | Bio-Rad Laboratories | 170-5060 |
| 5% milk powder | Carl Roth | T145.1 |
| RNasin Plus RNase inhibitor | Promega | PRN2615 |
| Lys-C | Wako Chemicals | 125-02543 |
| Trypsin | Promega | V5111 |
| Igepal CA-630 | Sigma-Aldrich | I8896 |
| Triton X-100 | Sigma-Aldrich | 9002-33-1 |
| Protein A Dynabeads | Thermo Fisher | 10002D |
| RNase I | Ambion | AM2295 |
| RNase A | Nacalal Tesque | 9001-99-4 |
| Proteus Clarification Mini Spin Column | Generon | GEN-MSF500 |
| ssRNA ligase | New England Biolabs | M0204S |
| PEG400 | Sigma-Aldrich | 202398 |
| sodium butyrate | Sigma-Aldrich | B5887-250MG |
| ascorbic acid | Sigma-Aldrich | A8960 |
| Hydrocortisone | Sigma-Aldrich | H0396 |
| Insulin | Thermo Fisher | 12585014 |
| Protan BA85 Nitrocellulose Membrane | Thermo Fisher | LC2009 |
| proteinase K | Roche | 3115828001 |
| Phase Lock Gel Heavy tube | 713-2536 | VWR |
| Glycoblue coprecipitant | Ambion | 9516 |
| Costar SpinX column | Corning Incorporated | 8161 |
| CircLigase ssDNA | Epicenter | CL4111K |
| Accuprime Supermix 1 | Thermo Fisher | Accuprime Supermix 1 |
| Agencourt AMPure XP beads | Beckman-Coulter | Agencourt AMPure XP beads |

**Critical Commercial Assays**

| assay | SOURCE | IDENTIFIER |
|-------|--------|------------|
| P3 Primary Cell 4D-Nucleofector Kit | Lonza | V4XP-3012 |
| Nucleofector Kit V | Lonza | VCA-1003 |
| RiboZero | Epicenter | Cat# MRZG12324 |
| TruSeq stranded total RNA Sample Prep Kit | Illumina | Cat# 20020599 |
| QuantSeq mRNA 3’ end sequencing kit | Lexogen | Cat# SKU 015.96 and SKU 016.96 |
| alkaline phosphatase staining | Sigma-Aldrich | AB0300-1KT |
| Agilent 2100 Bioanalyzer with RNA Pico 6000 kit | Agilent | 5067-1513 |
| Qubit dsDNA HS Assay Kit | Thermo Fisher | Q32854 |
| Inside Stain kit | Miltenyi Biotec | 130-090-477 |
| Nucleofector MEF 1 Kit | Lonza | VPD-1004 |
| Nascent RNA Capture Kit | Thermo Fisher | C10365 |

**Deposited Data**

| data | SOURCE | IDENTIFIER |
|------|--------|------------|
| QuantSeq and RNaseq libraries to study developmental alternative polyadenylation in hESC, its progeny and upon TDP-43 KD | This paper | E-MTAB-5090. Accessible via [https://www.ebi.ac.uk/arrayexpress/](https://www.ebi.ac.uk/arrayexpress/) |
| RNaseq libraries to study developmental alternative polyadenylation in human primitive streak progenitors derived from hESCs. | This paper | E-MTAB-5099. Accessible via [https://www.ebi.ac.uk/arrayexpress/](https://www.ebi.ac.uk/arrayexpress/) |

(Continued on next page)
Continued

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| QuantSeq libraries to study TDP-43 dependent alternative polyadenylation in mESCs and upon paraspeckles/TDP43 perturbations | This paper | E-MTAB-5091. Accessible via https://www.ebi.ac.uk/arrayexpress/ |
| RNaseq libraries to study TDP-43 dependent alternative polyadenylation in mESCa and cTdp43KO | This paper | E-MTAB-5097. Accessible via https://www.ebi.ac.uk/arrayexpress/ |
| iCLIP datasets for TDP-43 | This paper | E-MTAB-5100. Accessible via https://www.ebi.ac.uk/arrayexpress/ |
| Nascent RNaseq in conjunction with Illumina TRUseq method to sequence total RNAs including short lived RNAs using highly strand-specific next-generation sequencing (NGS) libraries | This paper | E-MTAB-5115 |
| QuantSeq libraries to study gene expression changes in Neat1/pA compared to parental undifferentiated hESCs | This paper | Accessible via https://expressrna.org/ |
| Fractionation RNaseq in conjunction with Illumina TRUseq method | This paper | E-MTAB-5116 |
| RNaseq differentiation of hESCs to endoderm, mesoderm, ectoderm, and hepatoblasts | Gifford et al., 2013 | GSM1112847, GSM1112845, GSM1112846, GSM1112844, GSM1112835, GSM1112833, GSM1112834, GSM1112837, GSM1124072, GSM1124072 |
| iCLIP datasets for TDP-43 in SH-SY5Y cells | Tollervey et al., 2011 | E-MTAB-527 Accessible via https://www.ebi.ac.uk/arrayexpress/ |
| Analysis of Pro-Inflammatory Gene Activation and RNA Processing by RNaseq of Nascent Transcripts I | | GSE32916 and GSE38892, accessible via GEO accession viewer |
| Mendeleev dataset; Additional data composed of Sanger sequencing and confirmatory gel electrophoresis confirming the CRISPR-Cas9 edited human and mouse embryonic stem cells. | This paper and Mendeleev data | https://data.mendeley.com/datasets/hbr8p7r526/2 |

**Experimental Models: Cell Lines**

| HEK293 Flag-WT TDP-43 tetracycline inducible | Budini et al., 2015 |
| HeLa | ATCC: CCL-2 RRID: CVCL_0045 |
| primary neonatal human dermal fibroblast (NHDF) | ATCC CRL-2522 |
| secondary human fibroblasts (HiF-Ts) | Cacchiarelli et al., 2015 |
| HAP1 | ATCC 204508 |
| hESCs H9 | Wicell WB66593 |
| V6.5 mESCs | ATCC SCRC-2011 |
| e14tg2a ICE | ATCC SCRC-1029 |
| cTdp-43 KO | Ling et al., 2015 |
| IDG3.2 | Hitz et al., 2007 (C57BL/6J x 129S6/SvEvTac)F1 mESC line |

**Experimental Models: Organisms/Strains**

| Mouse: CD-1 | Charles River |
| Mouse: mT/mG | Muzumdar et al., 2007 |
| Mouse: mT/mG | Mouse: mT/mG |

**Oligonucleotides**

| Stellaris smFISH Probe: Human NEAT1 Middle Segment with Quasar 570 | Stellaris | SMF-2037-1-BS |
| Stellaris smFISH Probe: Human NEAT1 5’ segment conjugated to Quasar®670 or Quasar®570 | Stellaris | SMF-2036-1 |
| Stellaris smFISH Probe: Mouse Neat1 middle segment conjugated to Quasar®670 | Stellaris | Customized product |
| Stellaris smFISH Probe: Mouse Neat1 5’ segment conjugated to Quasar®570 | Stellaris | Customized product |

(Continued on next page)
| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Stealth RNAi siRNA | Thermo Fisher Scientific | 12935-200 |
| TDP-43 siRNA, 5'-GGCUCAAGCAUGGAUUCUA- | Customized product | |
| TDP-43 siRNA, 5'-GGCUCAAGCAUGGAUUCUA- | Dharmacon | Customized product |
| GAPDH_F | GCTCATTTCCTGGTGATGACAAAGC |
| GAPDH_R | GAGATTAGTGCTGGTGCGGG |
| OCT4_F | CAATTTCAGAGCCCTCAAG |
| OCT4_R | AAACGGGAGATGTCGTT |
| MESP_F | CTGATCAGAGGAAGAAGCT |
| MESP_R | GCAGTCTGGCCAAAGGACCA |
| T_F | CAAATCTCAGAGGAGAAGC |
| T_R | ACAAAATTCTCAGAGGAGAAGC |
| MIXL_F | CCGAGTCAGAGGAGAAGC |
| MIXL_R | CTCTGACAGAGGCAGG |
| SOX17_F | GGCGACAGAGGAATCCAG |
| SOX17_R | CCACGTGGCCAGGC |
| Lefty1_F | TGACATCCAGGAGGAGG |
| Lefty1_R | AACTAAACAGGAGGAGG |
| EOMES_F | AGAGAGATTTCCAGG |
| EOMES_R | TTGTAAGACTATCATCGG |
| GDF3_F | GAGACTATGTTCTGAAGAG |
| GDF3_R | GAGTTAAGAGGAGAAGC |
| FOXD3_F | CTACTACAGAGGAGGAGG |
| FOXD3_R | GTGAGTGAGAGGGATG |
| Neat1.screen_F | GTGGGTGGTGAAGGAGG |
| Neat1.screen_R | CTTCCTCCAGAGAAGG |
| NEAT1pA.screen_F | TGAACAGAAGGAGGAGG |
| NEAT1pA.screen_R | CTTTCCTCAGAGGAGG |
| Neat1.screen_F | GTGGGTGGTGAAGGAGG |
| Neat1.screen_R | CTTCCTCCAGAGAAGG |
| Sox1 Δ2mir21.screen_F | TTAAAGGCGAAGGAGG |
| Sox1 Δ2mir21.screen_R | GCGAGTCAGAGGAGG |
| Neat1ΔTH.screen_F | TTACTACAGAGGAGG |
| Neat1ΔTH.screen_R | TCTTTCAGAGGAGG |
| Gapdh_F | TTCAACGAGGAGG |
| Gapdh_R | CCGCGTTGCTCCAGG |
| Neat1_1_F | TTGGGAGATTTCTGAAGG |
| Neat1_1_R | TCAAGTGGCGGAGG |
| Neat1_2_F | GATCGGAGCCAGGAGG |
| Neat1_2_R | AGCTTCAGAGGAGG |
| mmu-mir-21 | QIAGEN | MIMAT0000076: 5'UAG CULU AUC AGA CUG AUG UUG A, MIMAT0004494: 5'UUG CUC CAC CAG UCG UUG GU |
| TaqMan Gene Expression Assay NEAT1, isoform v2 | Thermo Fisher | Hs03924655_s1 |
| TaqMan Gene Expression Assay, Gene Symbol: SOX2 | Thermo Fisher | Hs01053049_s1 |
| TaqMan Gene Expression Assay, Gene Symbol: GAPDH | Thermo Fisher | Hs02758991_g1 |
| TaqMan Gene Expression Assay, Neat1, Gene Symbol NEAT1 isoform v1 | Thermo Fisher | Hs01008264_s1 |
| TaqMan Gene Expression Assay, Gene Symbol POU5F1 | Thermo Fisher | Hs0099632_g1 |

(Continued on next page)
| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| TaqMan® Gene Expression Assay, Gene Symbol | Thermo Fisher | Hs03309805_s1 |
| TaqMan® Gene Expression Assays, gene symbol Neat1v1 | Thermo Fisher | Mm03455878_s1 |
| **Recombinant DNA** | | |
| pmaxGFP plasmid with insertions of modified SOX2-UTR | Addgene | 16007 |
| PMIRH21PA-1-GVO-SBI | System Biosciences | PMIRH21PA-1-GVO-SBI |
| PX330-B/B | Addgene | 42230 |
| pcDNA6/TR | Addgene | 5676 |
| pCXLE-hMLN | Addgene | 27079 |
| MIP 247 CoMiP 4in1 with shRNA p53 | Addgene | 63726 |

| Software and Algorithms | | |
|-------------------------|--------|-----------|
| Bowtie2 | [http://bowtie-bio.sourceforge.net/bowtie2/index.shtml](http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) |
| TopHat2 | RRID: SCR_013035; [http://tophat.cbcb.umd.edu/](http://tophat.cbcb.umd.edu/) |
| STAR | Dobin et al., 2013 | RRID: SCR_015899; [https://github.com/alexdobin/STAR](https://github.com/alexdobin/STAR) |
| Cufflinks | | RRID: SCR_014597; [http://cole-trapnell-lab.github.io/cufflinks](http://cole-trapnell-lab.github.io/cufflinks) |
| expressRNA | Rot et al., 2017 | [http://www.expressrna.org/](http://www.expressrna.org/) |
| BLAST+2.3.0 | | RRID: SCR_001598; [https://blast.ncbi.nlm.nih.gov/Blast.cgi](https://blast.ncbi.nlm.nih.gov/Blast.cgi) |
| R | R Project for Statistical Computing | RRID: SCR_001905; [http://www.r-project.org/](http://www.r-project.org/) |
| DEXseq | [https://github.com/roryk/DEXSeq/](https://github.com/roryk/DEXSeq/) |
| Progenesis QI software | version 2.0, Nonlinear Dynamics | |
| Mascot (version 2.5.1) | Matrix Science | [http://www.matrixscience.com/whats_new/mascot-2-5-1-1-patch-release.html](http://www.matrixscience.com/whats_new/mascot-2-5-1-1-patch-release.html) |
| EdgeR | Robinson et al., 2010 | [https://bioconductor.org/packages/release/bioc/html/edgeR.html](https://bioconductor.org/packages/release/bioc/html/edgeR.html) |
| iCount | [http://icount.fri.uni-lj.si/](http://icount.fri.uni-lj.si/) |
| HTSeq | [https://htseq.readthedocs.io/en/release_0.11.1/](https://htseq.readthedocs.io/en/release_0.11.1/) |

| Other | | |
| NEATΔpA.gRNA.up.fwd | CACCAGTGCAAAACACATTACTGTGTTG |
| NEATΔpA.gRNA.up.rev | AAACACGACAGTAATGTTGTCGT |
| NEATΔpA.gRNA.down.fwd | CACCAGTGCAAAACACATTACTGTGTTG |
| NEATΔpA.gRNA.down.rev | AAACACGACAGTAATGTTGTCGT |
| Neat1ΔpA.gRNA.up.fwd | caccGGAAGCCTCTTTGAATTGTTG |
| Neat1ΔpA.gRNA.up.rev | caccGGAAGCCTCTTTGAATTGTTG |
| Neat1ΔpA.gRNA.down.fwd | caccGGAAGCCTCTTTGAATTGTTG |
| Neat1ΔpA.gRNA.down.rev | caccGGAAGCCTCTTTGAATTGTTG |
| Neat1ΔpA.GRNA.down.rev | caccGGAAGCCTCTTTGAATTGTTG |
| Neat1ΔpA.GRNA.down.rev | caccGGAAGCCTCTTTGAATTGTTG |
| Neat1ΔpA.GRNA.down.rev | caccGGAAGCCTCTTTGAATTGTTG |
| Sox2Δmir21_gRNA1_up | caccGAGTTATCATCAGGTTTTC |
| Sox2Δmir21_gRNA1_down | caccGAGTTATCATCAGGTTTTC |
CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by Micha Drukker (micha.drukker@helmholtz-muenchen.de).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture

All mouse ESC lines e.g., e14tg2a, IDG3.2 (Hitz et al., 2007) and tamoxifen inducible (concentration of 1μg/mL or 2.5μM of (Z)-4-Hydroxytamoxifen, H7904, Sigma-Aldrich) Tardbp conditional knockout (termed cTdp-43 KO (Ling et al., 2015)) were maintained on 0.1% gelatin-coated plates in 1:1 Neurobasal (21103049) DMEM (11320033) medium containing N2 (17502048) and B27 (17504044) supplements, 1% Glutamax (35050), 1% nonessential amino acids (1140050), 0.1 mM 2-mercaptoethanol (31350-010) (all Thermo Fisher Scientific), 1μM MEK inhibitor PD0325901 (1408, Axon Medchem), 3μM GSK3 inhibitor CHIR99021 (4953/50 Tocris), and 1,000 U/ml LIF (produced in-house), a condition named 2iLIF. Cells were passaged using 0.25% Trypsin-EDTA (25200056, Thermo Fisher Scientific). Spontaneous differentiation was induced by the removing pathway inhibitors and LIF and replacing N2B27 supplements with 10% HyClone Fetal Bovine Serum (FBS) (SV30160.03HI GE Healthcare). Fresh medium was applied daily.

HESCs (line H9) were maintained in mTeSR1 medium (05850, STEMCELL Technologies) on Matrigel-coated plates (356234, Corning) prepared by 1:100 dilution, and 5 mL coating of 10 cm plates for 1 h at 37°C. Colonies were passaged using the gentle cell dissociation reagent (07174, StemCell Technologies). HESCs were differentiated by Trypsin-EDTA, dissociation to single cells and seeding (100,000 cells/cm²) on Matrigel-coated plates in DMEM/F12 medium (11320), supplemented with 20% KnockOut Serum Replacement (KSR; 10828028, Thermo Fisher Scientific), Glutamax, nonessential amino acids, and β-mercaptoethanol, or B27 supplemented as above with either 10 μM CHIR99021, 250ng/ml recombinant WNT3A (kindly provided by Derk ten Berge), 0.5 μM RA (R2625, Sigma-Aldrich) or 50ng/ml BMP4 (314-BP; R&D Systems). Fresh medium was applied daily.

HEK293T Flag-WT TDP-43 tetracycline inducible cell line was cultured as previously described (Budini et al., 2015). These HEK cells were transfected using Lipofectamine 3000 (L3000001, Thermo Fisher Scientific) according to manufacturer’s instructions. Fresh medium was applied daily. All cells were grown at 37°C in 5% CO2.

 Generation of h/mESC lines harboring edited NEAT1/Neat1

The NEAT1ΔpA hESC line was generated by deleting a 500 nt region surrounding the proximal pA site including the regulatory cis-elements (Naganuma et al., 2012). Based on Raj et al. (2008) forward and reverse gRNAs (listed below) with BbsI restriction site overhangs were designed, phosphorylated, annealed and cloned into BbsI digested Cas9-2A-GFP vector (48138 Addgene). One million cells were transfected with 5μg plasmid pairs using the P3 Primary Cell 4D-Nucleofector Kit (V4XP-3012 Lonza) and the 4D-Nucleofector Platform (Lonza), program CB-156. 10-14 days later clones were picked, expanded and PCR-screened for identifying clones with alleles harboring deletions.

A similar approach was used for deleting 60 nt surrounding the proximal pA site in Neat1, the region corresponding to the triple helix (∇TH) in Neat1_2, and miR-21 binding site in Sox2. 5 × 10^5 mESCs were plated on 0.2% gelatin-coated wells of a 6 well plate, and were transfected 3 h later (Lipofectamine 3000, L3000008, Thermo Fisher Scientific) with 2 μg total plasmid DNA made of a pair of SpCas9-2A-Puro plasmids (Plasmid #62988, Addgene), each containing single guide RNAs targeting the respective forward or reverse region in Neat1 (listed below). 48 h later the cells were reseeded and treated with media supplemented with 1 μg/ml puromycin (P8833, Sigma-Aldrich). 5-6 days later clones were picked, expanded in 96-well plates and PCR-screened for identifying clones with alleles harboring deletions. Sequences of gRNAs and primers used for generating and screening genetically edited h/mESCs are listed in the STAR Methods section.

 Generation of HAP-1 cells harboring edited NEAT1

PX330-B/B containing two sgRNAs (2 μg) and pcDNA6/TR (0.2 μg) containing the blasticidin resistance gene (Thermo Fisher Scientific) were co-transfected into HAP1 cells (1.5 × 10^6 cells) by Nucleofector Kit V (VCA-1003, Lonza) with a Nucleofector device (Lonza)

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Sox2Δmir21_gRNA2_up | caccgATTTAGGACCGTTACAACAA | aaacTGTGGTGAACCTGCCTAAATc |
| Sox2Δmir21_gRNA2_down | aaacTGTGGTGAACCTGCCTAAATc |
| Neat1ΔTH_gRNA1_fwd | caccgaaggaacgcgtactgca |
| Neat1ΔTH_gRNA1_rev | AAACtgcaaggaacgcgtactgca |
| Neat1ΔTH_gRNA2_fwd | caccgaaggaacgcgtactgca |
| Neat1ΔTH_gRNA2_rev | aaacCAGCGGTGTTTTTCTCTC |

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Sox2Δmir21_gRNA2_up | caccgATTTAGGACCGTTACAACAA | aaacTGTGGTGAACCTGCCTAAATc |
| Sox2Δmir21_gRNA2_down | aaacTGTGGTGAACCTGCCTAAATc |
| Neat1ΔTH_gRNA1_fwd | caccgaaggaacgcgtactgca |
| Neat1ΔTH_gRNA1_rev | AAACtgcaaggaacgcgtactgca |
| Neat1ΔTH_gRNA2_fwd | caccgaaggaacgcgtactgca |
| Neat1ΔTH_gRNA2_rev | aaacCAGCGGTGTTTTTCTCTC |
using program “X-005” according to the manufacturer’s instructions. To enrich the plasmid-transfected cells, the HAP1 cells were treated with 20 μg/ml blasticidin (ant-bl-05, InvivoGen) for 3 days, starting 1 day after transfection. Subsequently, the cells were diluted into 96-well plates for creating clones. Clones were lysed by proteinase K (200 μg/ml, 3115828001 Roche, 20 mM Tris-HCl pH 8.0, 5 mM EDTA, 400 mM NaCl, and 0.3% SDS at 55 C for 1 h, followed by proteinase K inactivation (95 C for 15 min). Then, the lysates were subjected to PCR analysis to amplify the genomic regions flanking the guide RNA target sites for detecting deletions or insertions using KOD FX Neo enzyme (KFX-201, TOYOBO). The indel-positive clones were further confirmed by sequencing.

Specifically, to establish HAP1 NEAT1 ΔUG cell line lacking three NEAT1_2 domains (5942-7025 nt, 9201-11091 nt, and 21061-22012 nt), these domains were deleted by sequential CRISPR/Cas9-mediated deletions. First, NEAT1_2 5942-7025 nt (5.9-7 kb) domain was deleted by sgRNAs (GGCGGGTCGTCCTAACTAGC and CATTTAAACCTTCTTCCCCG) in HAP1 WT cells. Second, NEAT1_2 21061-22012 nt (21-22 kb) was deleted by introducing two sgRNAs (TACGCGGGGAACGTGCCAC and TCCGACTTCATTTCAAGATGA) into HAP1 Δ5.9-7 kb cells. Third, NEAT1_2 9201-11091 nt was deleted by introducing two sgRNAs (GTACCTTAAACAGGCGGGG and TATTACCTTGCGCTAGGGG) into HAP1 Δ5.9-7 kb/21-22 kb cells to establish NEAT1 ΔUG cell lines. To establish HAP1 NEAT1 ΔUG/UGx60 cell line by knocking-in UG repeats (UG x 60) into the HAP1 NEAT1 ΔUG cell line, PX330-B/B (2 ug) containing two sgRNAs targeting to knock-in vector (GCATCGTACGCGTGTT) and NEAT1 genomic region (NEAT1_2 22.1 kb region) (GCGGGCTTCTCGTGACCT) were co-transfected with knock-in vector (pcDNA3) containing 60 TG repeats (generated by Genscript) and pcDNA6/TR plasmids (0.2 ug) into HAP1 NEAT1 ΔUG cells as described above. The insertion was detected by PCR and further confirmed by sequencing.

Overexpression of TDP-43

A mESC line (parental e14tg2a, kindly provided by Michael Kyba) harboring doxycycline inducible iTPD-43-eGFP transgene was generated by targeting the Rosa26 locus of the A2lox.Cre line with the p2lox (plasmid #43653, Addgene) construct cloned with TDP-43-eGFP fusion cassette. The recombination was performed as previously described (Kyba et al., 2002), and 24 h after transfection cells were selected with 250 ng/ml G418 (InvivoGen, ant-gn-1) for 7 days. Clones were picked, expanded and validated. This line was subsequently modified by deleting the 100 nt corresponding to the UUG-rich TPD-43 binding site in Neat1. Constitutive overexpression of TPD-43 during human fibroblast reprogramming was mediated by transduction or transfection of myc-tagged TPD-43 (Schwenk et al., 2016). Empty vector was used as a control.

Knockdown of TDP-43

2x10^5 Accutase-treated single hESCs were seeded 24 h prior to transfection on Matrigel-coated 6-well plates with mTeSR1 medium supplemented with ROCK inhibitor. Per well, siRNA transfection was performed using 5 μM siRNA duplexes mixed with 5 μL Lipofectamine RNAiMAX (13778150, Thermo Fisher Scientific) was used as the negative control. Transduction was used for knocking-down of TDP-43 via shRNAs for global analysis of PASs in parallel to the use of a luciferase shRNA vector as a control (Schwenk et al., 2016). Empty vector was used as a control.

Plasmids

The 3’ UTR of SOX2 was ectopically express by cloning the complete UTR (1263 nt including 100 nt downstream of distal pA site) immediately after the coding sequence of eGFP in the pmaxGFP plasmid (#16007 addgene), producing the pmaxGFP-SOX2-UTR construct. Two variants of this plasmid were prepared by deleting the mir-21 binding site (AAATGTCCATTGTTTATAAGCTGA) producing pmaxGFP-SOX2-UTRΔmiR21, or the regulatory sequence spanning the proximal pA site (GGAAATGGGAGGGGTGCAAAA CAGAGGAGTAGAAGAACGATGAGAAAAACCCGTCAGCTCAAAAAAGAAAAAGGAAAAAATCCCATCA) producing pmaxGFP-SOX2UTRΔpA. To ectopically express mir-21, the GFP sequence of plasmid PMIRH21PA-1-GVO-SBI (System Biosciences), which codes for the microRNA precursor of mir-21, was replaced by the sequence of the gene encoding tdTomato fluorescent protein.

Reprogramming of human fibroblasts

1.5x10^6 primary neonatal human dermal fibroblast (NHDF) (ATCC, CRL-2522) were transfected using the Nucleofector MEF 1 Kit (VPD-1004, Lonza) with 6 μg of plasmid MIP 247 CoMIP 4in1 and 3 μg of pCXLE-hMLN (63726 and 27079, Addgene). Cells were pulsed with T-020 or N-024 program using Nucleofector 2b (AAB-1001, Lonza), and were plated on Matrigel-coated plates with fibroblast medium composed of DMEM high glucose (11965092, Thermo Fisher Scientific) and 10% HyClone Fetal Bovine Serum supplemented with ROCK inhibitor. Per well, siRNA transfection was performed using 5 μM siRNA duplexes mixed with 5 μL Lipofectamine RNAiMAX (13778150, Thermo Fisher Scientific) and 0.1mM sodium butyrate, 0.1mM Hydrocortisone (H0396, Sigma Aldrich) and 64 μg/mL ascorbic acid. Between days 10-15 first iPSC-like colonies appeared, after medium was changed to Essential 8 medium (E7 + 2ng/ml TGFbeta1) (100-21, Peprotech) supplemented with 0.1 mM sodium butyrate, 0.1mM Hydrocortisone (H0396, Sigma Aldrich) and 64 μg/mL ascorbic acid. Between days 10-15 first iPSC-like colonies appeared, after medium was changed to Essential 8 medium (E7 + 2ng/ml TGFbeta1) (100-21, Peprotech) supplemented with 0.1 mM sodium butyrate, 0.1mM Hydrocortisone (H0396, Sigma Aldrich) and 64 μg/mL ascorbic acid. Between days 10-15 first iPSC-like colonies appeared, after medium was changed to Essential 8 medium (E7 + 2ng/ml TGFbeta1) (100-21, Peprotech) supplemented with 0.1 mM sodium butyrate, 0.1mM Hydrocortisone (H0396, Sigma Aldrich) and 64 μg/mL ascorbic acid in low O2 (5% O2) conditions. For
reprogramming of secondary human fibroblasts (HiF-Ts), cells were cultured and reprogrammed as described (Cacchiarelli et al., 2015). The number of reprogrammed colonies was analyzed by alkaline phosphatase staining (AB0300-1KT, Sigma Aldrich).

**ANIMALS**

**Generation of chimeras**

Tetraploid chimeras were generated according to standard protocols (Engert et al., 2013). Embryos were collected from the mT/mG expressing mouse line (Muzumdar et al., 2007), maintained on C57/Bl6J background.

**Animal data**

Mouse keeping was done at the central facilities at HMGU in accordance with the German animal welfare legislation and acknowledged guidelines of the Society of Laboratory Animals (GV-SOLAS) and of the Federation of Laboratory Animal Science Associations (FELASA).

**Immunofluorescence of embryos**

Immunofluorescence whole-mount staining was performed in the following way. Briefly, embryos were isolated, fixed for 20 minutes using 2% PFA in PBS, permeabilized using 0.1% Triton X-100 in 0.1 M glycine pH 8.0. After blocking using 10% FCS, 3% donkey serum, 0.1% BSA, 0.1% Tween 20 for 2 h, embryos were incubated with the primary antibody o/n at 4°C in blocking solution. After several washes in PBS containing 0.1% Tween-20 (PBST) embryos were incubated with secondary antibodies (donkey anti-goat 488, donkey anti-rabbit 555 each 1:800) in blocking solution for 3 h. During the final washes with PBST, embryos were stained with 4′,6-diamidino-2-phenylindole, dihydrochloride (DAPI), transferred into 15% and 30% glycerol and embedded between two coverslips using 120µm Secure-Seal spacers (Invitrogen, S24737) and ProLong Gold antifade reagent (Invitrogen, P36930). Antibodies: Foxa2 1:1000 (Santa Cruz, sc-6554), Brachyury 1:500 (N-19, Santa Cruz, sc17743), RFP 1:1000 (Rockland, 600-401-379S).

**METHOD DETAILS**

**RNA preparation and qRT-PCR assays**

Total RNA was prepared from cell pellets using miRNeasy Micro Kit (217084, QIAGEN) according to the manufacturer’s instructions. First-strand cDNA synthesis of 1 µg total RNA was performed using SuperScript III Reverse Transcriptase (18080085, Invitrogen), according to the manufacturer’s guidelines. qRT-PCR reactions were performed using Power SYBR Green Master Mix (4367659, Thermo Fisher Scientific). miR-21 was analyzed using the miScript protocol (218193, QIAGEN). Primers used in this study are listed in the STAR Methods section.

**RNA sequencing**

3 µg total RNA was treated with TURBO DNase (am2238, Thermo Fisher Scientific) according to manufacturer’s instructions followed by RNaseq analysis (74204, QIAGEN). Microcapillary electrophoresis on Agilent 2100 Bioanalyzer with RNA Pico 6000 kit (5067-1504, Agilent) was used to analyze RNA quality (RIN values > 8). For total RNA sequencing and per library, 1 µg RNA was depleted of rRNA using RiboZero Gold (Human/Mouse/Rat) kit followed by a cleanup step and library preparation using TruSeq kit (5067-1513, Agilent) was used to perform the RNA-Seq experiment.

**Western blotting**

Cells were trypsinized and lysed using RIPA buffer, containing phosphatase (Sigma-Aldrich, 4906837001) and protease (Merck, 539134) inhibitors. After addition of 2x SDS loading buffer with 2-Mercaptoethanol (Sigma-Aldrich, M3148) samples were heated to 95°C for 5 min. Samples were ran on Mini-PROTEAN TGX Stain Free Gels, 4%–15% (Bio-Rad Laboratories, 456-8086), and blotted using the Mini Trans-Blot Cell (Bio-Rad Laboratories, 1703930). Following 3 x 5 min washes with TBS-T, membranes were blocked with 5% milk powder (T145.1, Carl Roth) in TBS-T. Membranes were then incubated o/n at 4°C with 5% milk powder in TBS-T containing the primary antibody. After 3 x 5 min TBS-T washes, membrane was incubated with goat anti-rabbit IgG-HRP (sc-2030, Santa Cruz) or Goat anti-mouse IgG HRP-conjugated (115-035-003 Dianova) in 5% milk powder in TBS-T. Following 4 15 min washes
with TBS-T the membrane was incubated for 1 min with Clarity Western ECL Substrate (170-5060, Bio-Rad Laboratories) and imaged with ChemiDoc MP System (Bio-Rad Laboratories) or exposed to X-ray film. Antibodies used are listed in the STAR Methods.

**Immunofluorescence**

Cells were grown in Matrigel coated 8 well chamber slides (80826, Ibidi) and fixed with 4% PFA/DPBS solution (Thermo Fisher Scientific 16% Formaldehyde (w/v), Methanol-free, 28906) for 15 min at RT and permeabilized using 0.2% Triton X-100/DPBS solution for 15 min at RT. Primary and secondary antibodies were diluted per manufacturer recommended concentrations in 10%FBS/0.2% Triton X-100/DPBS and incubated respectively o/n at 4°C and 1 h at RT. The samples were washed with DAPI (50ug/ml) solution and imaged using a Zeiss Axiovert 200M epifluorescent microscope. Antibodies used are listed in the STAR Methods.

**Flow cytometry**

Plates were washed with PBS, and cells were dissociated by Trypsin-EDTA, followed by resuspension in 2% FBS 1mM EDTA PBS buffer, and incubation with antibodies for 30-60 min on ice. Cells were centrifuged, resuspended in buffer containing SYTOX blue for dead cell exclusion and analyzed using a FACs Aria III (BD Biosciences). Cell debris were excluded by forward and side scatter gating. FlowJo was used for data analysis.

For intracellular flow cytometry mESCs were dissociated by Accutase (A6964, Sigma-Aldrich), centrifuged and resuspended in 2% of methanol-free formaldehyde (28906, Thermo Fisher Scientific) for 10 min in RT. Inside Stain kit (130-090-477, Miltenyi Biotec) was used according to manufacturer’s protocol. Antibodies used are listed in the STAR Methods.

**Subcellular fractionation of hESCs**

5-10^6 cells were harvested, washed with ice-cold PBS and centrifuged at 500g 4°C. Cell pellets were gently resuspended in 380 µL cold cytoplasmic lysis buffer (50 mM Tris-pH 6.5, 100 mM NaCl, 300 mM Sucrose, 3 mM MgCl₂, 0.15% NP40) supplemented with 100 U RNasin Plus RNase inhibitor (PRN2615, Promega). After incubation on ice for 10 min, cells were briefly vortexed and centrifuged at 1000g 4°C for 3 min. The supernatant containing the cytoplasmic fraction was transferred, and centrifuged again at 4°C 5000g for 2 min to remove cell debris. Immediately after centrifugation 1 mL of RNA precipitation buffer (RPS; 9.5 mL 100% ETOH with 0.5 mL 3 M Sodiumacetate) was added and the supernatant was incubated at −20°C for 3 to 5 hours until further RNA purification.

The remaining pellet was washed three times with 400 µL cytoplasmic lysis buffer supplemented with an increasing concentrations 50 mM, 200 mM and 500 mM of Ammonium sulfate (for disrupting the endoplasmic reticulum attached to the nucleus but without breaking the nuclear envelope). For each washing step, the cell suspension was centrifuged at 4°C and 5000g for 2 min and the supernatant was processed for RNA extraction.

380 µL of cold Modified Wuarin-Schibler (MWS) buffer (10 mM Tris-HCl pH 7.0, 4 mM EDTA, 0.3 M NaCl, 1 M Urea, 1% NP-40) supplemented with 100 U RNasin Plus RNase inhibitor was added to the remaining pellet. Samples were vortexed for 30 s, incubated on ice for 5 min, then vortexed again and kept on ice for 10 min. The suspension was centrifuged at 1000g 4°C for 3 min and the resulting supernatant representing the nucleoplasmic fraction was processed for RNA extraction.

The pellet was washed three times by adding 800 µL of MWS buffer, vortexing for 30 s and centrifuging at 500g 4°C for 2 min. 1 mL QIAzol (79306, QIAGEN) was added to the remaining chromatin pellet and after short vortexing, the suspension was stored at −20°C until further usage.

The RNA fraction in RPS buffer was vortexed for 30 s and after adding 1 µL Glycoblue coprecipitant (Life Technologies, AM9516) centrifuged at 18000 g 4°C for 15 min. 1 mL QIAzol was added to the partially air-dried pellet. 10 µL of 0.5 M EDTA was added to all samples and heated up to 65°C and incubated for 10 min to resuspend the pellet. After cooling down, 200 µL of chloroform / isoamyl alcohol (24:1) was added, the solution was vortexed for 30 s and then centrifuged at 18000 g RT for 10 min. The upper aqueous phase was transferred into a new tube and the same volume of isopropanol together with 1 µL of Glycoblue was added. After o/n incubation at −20°C, the samples were centrifuged at 18000 g RT. The pellets were washed once by adding 1 mL 70% EtOH and centrifuging at 18000 g RT for 5 min. Pellets were air-dried for 10 min and resuspended in RNase-free water. Cytoplasmic, nucleoplasmic and chromatin fractions were analyzed by RNA-sequencing as outlined above. Random primed, strand specific cDNA libraries were prepared following Illumina TruSeq total RNA protocol (above). 81 – 92% of obtained reads aligned to the human genome allowing a single mismatch.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

**Single molecule (sm)FISH and paraspeckle quantification**

Based on Raj et al. (2008), h/mESCs that were grown on Matrigel coated sterile 2-well µ-Slides (80286, Ibidi) were washed twice for 5 min with 1x PBS, and fixed using 4% methanol-free formaldehyde (10321714, Thermo Fisher Scientific) for 10 min at RT. Following additional two wash steps, cells were permeabilized using 70% ethanol for 12 h at 4°C, and were washed twice again. Then preparations were incubated for 15 min with hybridization buffer prepared using 2x saline-sodium citrate (SSC) solution / 10% deionized formamide (4610-OP, Calbiochem). Hybridization with Stellaris FISH probes was done in a total volume of 50 µl hybridization buffer containing 50 µg competitor tRNA from E.coli (10109541001, Roche Diagnostics), 10% Dextran Sulfate (9011-18-1, VWR), 2 mg/ml
UltraPure BSA (AM2616, Thermo Fisher Scientific), and 10 mM vanadyl-ribonucleoside complex (S1402S, New England BioLabs) with probes at final concentration of 1 ng/μl. Preparations were covered with paraffin and incubated at 37°C for 5 h, and afterward washed twice with pre-warmed 2xSSC/10% formamide for 30 min at 37°C. Finally the preparations were washed twice with 1x PBS at RT, and then mounted using 10 μl ProLong Gold Antifade Reagent containing DAPI (9071S, New England BioLabs). The slides were imaged when the mounting medium was fully cured >12 h.

Probes were designed using the Probe Designer software from Biosearch Technologies and were provided by same vendor. Probes included were hNEAT1 middle segment conjugated to Quasar570 (SMF-2037-1), human NEAT1 5’ segment conjugated to Quasar670 or Quasar570 (SMF-2036-1), mouse NEAT1 middle segment conjugated to Quasar670 and mNEAT1 5’ segment conjugated to Quasar570 (sequences of probes available upon request). Quantification of hybridization signal was performed using spot detection algorithm Airlocalize (Trcek et al., 2017).

RNA-FISH and immunofluorescence
RNA-FISH - immunofluorescence was performed as previously described (Kawaguchi et al., 2015; West et al., 2016). The RNA probes were synthesized using T7 or SP6 RNA polymerase and a digoxigenin (DIG) RNA labeling kit (11175025910, Roche). Linearized plasmids (1 μg) containing a NEAT1 fragment (NEAT1 1-1000 nt) were used as templates. Cells were grown on coverslips (Matsunami; micro cover glass; 18 mm round; thickness, 0.16–0.19 mm) and treated with 5 uM proteasome inhibitor (MG132, Sigma-Aldrich) for 6 h to enhance NEAT1 expression (Hirose et al., 2014). Then cells were fixed with 4% paraformaldehyde/PBS at room temperature for 10 min, washed using 1 x PBS, permeabilized with 0.5% Triton X-100 (9002-93-1, Sigma-Aldrich)/PBS for 5 min, and washed three times with 1 x PBS. Subsequently, the cells were dipped in 100% ethanol for 5 min and then dried. Dehydrated coverslips were incubated with pre-hybridization solution (50% formamide, 1 x Denhardt’s salt, 2 x SSC, 100 mM EDTA, 100 ng/ml yeast tRNA, and 0.01% Tween 20) at 55°C for 1 h and then incubated with hybridization solution (as above with 5% Dextran sulfate) containing the DIG-labeled RNA probes (final concentration: 100 ng/cover slip) at 55°C o/n. After hybridization, the coverslips were washed twice with pre-warmed wash buffer (50% formamide, 2 x SSC, and 0.1% Tween 20) at 55°C for 15 min. Then, excess RNA probes were digested with 10 μg/ml RNase A (9001-99-4, Nacalai Tesque) in NTET buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA, 500 mM NaCl, and 0.1% Tween 20) at 37°C for 30 min. The coverslips were washed with buffer (2 x SSC and 0.01% Tween 20) at 55°C for 15 min and washed twice with a second buffer (0.1 x SSC and 0.01% Tween 20) at 55°C for 15 min. The coverslips were subsequently washed with TBST (1 x TBS containing 0.1% Tween 20) and incubated with 1 x blocking solution (Blocking reagent [Roche] diluted with TBST) for blocking at RT for 1 h. Then, the coverslips were incubated with primary antibodies, Anti-Digoxigenin mouse monoclonal antibody (clone 21H8, Abcam), Anti-TDP-43 rabbit polyclonal antibody (Proteinitech, 1:100), anti-PSPC1 rabbit polyclonal antibody (1:1000; Naganuma et al., 2012) in 1 x blocking solution at room temperature for 1 h, washed three times with TBST for 5 min, and incubated with secondary antibodies, anti-mouse IgG Alexa 488, and anti-rabbit IgG, Alexa-568 (both Thermo Fisher Scientific) in 1 x blocking solution at RT for 30 min, and washed three times with TBST for 5 min. The coverslips were mounted with VECTASHIELD Hard Set Mounting Medium with DAPI (Vector). Confocal images were acquired using a confocal laser scanning microscope FV1000D (Olympus). Cells were automatically counted using Fiji software. First, nuclei or paraspeckles were segmented using either DAPI or FISH staining to identify areas of interest. Then gaussian blur (1.2) was applied and threshold for each object was calculated using Otsu’s method. Mean intensity values or correlation coefficients were measured for the respective areas.

iCLIP protocol and its analysis
The iCLIP protocol was performed as described previously (Huppertz et al., 2014) with several modifications. Cells were irradiated by UV once with 160 mJ/cm² using Stratalinker 1800 at 254 nm. TDP-43 was immunoprecipitated with protein A Dynabeads (10002D, Invitrogen) conjugated to rabbit-anti TDP-43 (Sigma-Aldrich, SAB4200006) or GFP (Life, Technologies, A6455). The region corresponding to 55–100kDa complexes was excised from the membrane to isolate the RNA, and sequenced using Illumina HiSeq, generating 50-nucleotide single-end reads, sequencing depth was 15-20 Mio reads per library. Analysis of reproducibility of cross-linked sites, identification of the significant iCLIP crosslinked clusters and z-score analysis of enriched pentamers was performed as described (Tollervey et al., 2011) and data was processed by the iCount software (https://github.com/tomazc/iCount).

RNA-Seq differential expression
To identify differentially expressed genes from RNA-Seq data, sequencing reads were mapped using TopHat2 (Trapnell et al., 2009), and gene count tables were produced using HTSeq and Ensembl genome version 74. Finally, edgeR (Robinson et al., 2010) was applied to analyze P value and fold-expression changes.

Identification of polyA sites
QuantSeq read data were aligned to hg19 reference genome with STAR (Dobin et al., 2013). Reads that did not align uniquely were filtered out. Identification of polyA-sites with QuantSeq relies on annealing of poly(T) primers to the polyA tails of mRNAs. To exclude cases of internal priming, where primer anneal to genomic polyA sequences, A-rich regions in vicinity of putative polyA sites were removed ([-10..10]). Additionally, only polyA-sites that were more than 125 nt apart from each other were treated separately, because
cleavage is not a nucleotide exact process, and therefore closely spaced polyA-site were pooled. PolyA site classification was done by including those sites that were adjacent to polyA-site signals [-100, 25] nt.

polyA sites were categorized based on presence of a preceding polyA signal using polyAR (Akhtar et al., 2010). Four classes were annotated, strong, weak, lacking polyA signal and non-categorized sites. The nucleotide composition, overlap with a published data-set of polyA-sites (Derti et al., 2012) and efficiency of cleavage (cDNA counts) for each class of sites was analyzed, and confirmed that strong and weak sites were the most reliable sites in terms of known nucleotide composition around polyA-sites.

Strong and weak polyA-sites were used to define position-dependent polyA-site regulation by TDP-43. Of a total 28,930 polyA sites (strong + weak), 23,309 sites were annotated to genes (Ensembl version 74). 17,286 polyA-site exhibited strong signals across experiments (a minimum of 10 reads needed to be detected in either test or control sample). Summing the counts across all experiments was used as an additional filtering step to avoid sites that resulted from inefficient cleavage. Only polyA sites that exhibited > 5% the level of another major site in the same gene were included. Since alternative long 3' UTRs are not yet fully annotated, 5k of the intergenic region downstream of each gene were included in this analysis (however only to the middle of the downstream gene if distance < 5kb). This led to the identification of 16,065 sites in 10,661 genes presented in (Table S5).

### Analysis of alternative polyadenylation (APA)

Genes included were those exhibiting two major pA-site (highly expressed in both test and control). Next, pA-site pairs were classified as same-exon, composite-exon or skipped-exon. For same-exon sites, the major site was compared to the sum of all other sites within the same exon (exon level). For composite-exon and skipped-exon pairs, the major site was compared to the expression of the other sites in the gene.

To estimate the level of change in expression between control and test groups, we calculated the “percent change” (pc) score, as follows:

\[
p_{\text{pc}} = \frac{\text{control}_{\text{PROXIMAL}}}{\text{control}_{\text{DISTAL}}} - \frac{\text{test}_{\text{PROXIMAL}}}{\text{test}_{\text{DISTAL}}} \quad [-1, 1]
\]

Positive values determine a higher ratio of control versus test in proximal versus distal sites, and the negative values represent the opposite trend. Fisher’s exact test was then used to determine the significance of the change. Genes were regarded as displaying significant changes in polyA site usage between test and control if Fisher’s p value was < 0.1. Lastly genes were classified as repressed (pc < -0.1), enhanced (pc > 0.1) and controls (abs(pc) < 0.1 and p value > 0.1). GO-term analysis of TDP-43 regulated alternatively polyadenylated transcripts was performed based on Blake et al. (2015).

### Visualizing position-dependent polyA site regulation using RNA-maps

After identifying genes exhibiting APA, summed crosslinked sites derived from iCLIP were plotted around the pA sites of repressed, enhanced and control genes (Figures S5E–S5G). This RNA-map approach is similar to plotting RNA-maps around splice sites in alternative-splicing context (Ule et al., 2006; Witten and Ule, 2011).

### mRNA/RBP occupancy (mRNA interactome capture)

Protocol was based on Castello et al. (2013). ~50X10^4-6 undifferentiated cells in 3 replicates of WT hESC and NEAT1ΔpA hESC were irradiated as described for iCLIP.

### FASP digest

Each 10 μg of RBPome were digested with a modified FASP procedure (Wiśniewski et al., 2009). Briefly, the proteins were reduced and alkylated using dithiothreitol and iodoacetamide, diluted with one volume of UA buffer (8 M urea in 0.1M Tris/HCl pH 8.5) and then centrifuged through a 30 kDa cut-off filter device (PALL, Port Washington). Samples were washed twice with UA buffer and twice with 50 mM ammonium bicarbonate prior to digest of the immobilized proteins on the filter for 2 h at RT using 1 μg Lys-C (Wako Chemicals) and for 16 h at 37°C using 2 μg trypsin (Promega). Tryptic peptides were collected by centrifugation (10 min at 14,000 g), and the samples were acidified with 0.5% TFA and stored at −20°C.

### Mass spectrometry

Before loading, the samples were centrifuged for 5 min at 4°C. LC-MS/MS analysis was performed on a QExactive HF mass spectrometer (Thermo Scientific) online coupled to an Ultimate 3000 nano-RSLC (Thermo Scientific). Approximately 0.5 μg of every digested sample was automatically injected and loaded onto the trap column at a flow rate of 30 μl/min in 3% ACN/ 0.1% FA. After 5 min, the peptides were eluted from the trap column and separated on the C18 analytical column (75 μm i.d. x 25 cm, Acclaim PepMap100 C18, 2 μm, 100A, Dionex) by a 90 min gradient from 5 to 25% ACN in 0.1% FA at 300 μl/min flow rate followed by a 5 min gradient from 25% to 40% ACN in 0.1% FA. Between each sample, the column was washed with 85% ACN for 5 min followed by equilibration at 3% ACN in 0.1% FA for 18 min. MS spectra were recorded at a resolution of 60000 with an AGC target of 3e6 and a maximum injection time of 50 ms from 300 to 1500 m/z. From the MS scan, the 10 most abundant peptide ions were selected for fragmentation via HCD with a normalized collision energy of 27, an isolation window of 1.6 m/z and a dynamic exclusion of 30 s. MS/MS spectra were recorded at a resolution of 15000 with an AGC target of 1e5 and a maximum injection time of 50 ms. Intensity threshold was set to 1e4 and unassigned charges and charges of +1 and > 8 were excluded.
Analysis of mRNA-interactome

Progenesis QI software (version 2.0, Nonlinear Dynamics) was used for label free quantification as described previously (Merl et al., 2012). Briefly, profile data of the MS and MS/MS scans were transformed to peak lists with respective peak m/z values, intensities, abundances (areas under the peaks) and m/z width. After reference selection, the retention times of the other samples were aligned by automatic alignment to a maximal overlay of all features. After exclusion of all features with only one charge or more than seven charges, all remaining MS/MS spectra were exported as Mascot generic file and used for peptide identification with Mascot (version 2.5.1) in the Ensembl Human protein database. Search parameters used were: 10 ppm peptide mass tolerance and 20 mmu fragment mass tolerance, one missed cleavage allowed, carboxymethylation was set as fixed modification, methionine oxidation and asparagine or glutamine deamidation were allowed as variable modifications. A Mascot-integrated decoy database search calculated an average false discovery of 1% when searches were performed with a mascot percolator score cut-off of 13 and an appropriate significance threshold p. Peptide assignments were re-imported into the Progenesis QI software and the abundances of all peptides allocated to each protein were summed up. Resulting normalized protein abundances were used further and compared to existing mRNA-interactome datasets obtained from HeLa (Castello et al., 2012) and HEK293 cells (Baltz et al., 2012). As the aim of this study was not to expand the repertoire of RBPs, but rather to determine the dynamics of high confidence RBPs bound to mRNAs, we used only the overlap between published and our datasets (388 high confidence RBPs) for analysis of mRNA interactome changes upon gain of paraspeckles in NEAT1 DpA cells compared to the parental line.

DATA AND SOFTWARE AVAILABILITY

Raw sequencing files (50 nt, single-end) have been deposited in the ArrayExpress archive (accessible at E-MTAB-5090; E-MTAB-5091; E-MTAB-5097; E-MTAB-5099; E-MTAB-5100; E-MTAB-5114, E-MTAB-5115, E-MTAB-5116) and https://expressrna.org/ platform that enables visualization and accession of the sequencing data. Similarly, iCLIP data are deposited and available also in https://imaps.genialis.com/ and validation of CRISPR-Cas9 modified cell lines are deposited to Mendeley data (https://data.mendeley.com/datasets/hbr8p7n526/2).