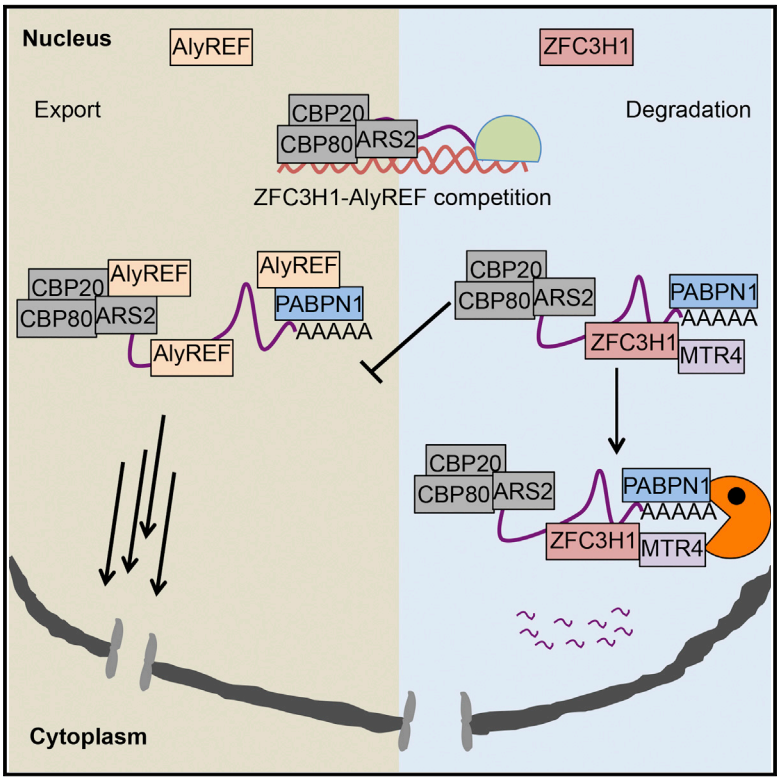
Article

Cell Reports

The RNA Exosome Adaptor ZFC3H1 Functionally Competes with Nuclear Export Activity to Retain Target Transcripts

Graphical Abstract



Highlights

- Abolished RNA exosome function leads to pA$^+$ RNA accumulation in nuclear foci
- pA$^+$ RNA foci are enriched with various transcripts and exosome adaptor proteins
- The exosome adaptor protein ZFC3H1 is required for pA$^+$ RNA foci formation
- ZFC3H1 functionally counteracts the mRNA export factor AlyREF

Authors

Toomas Silla, Evdoxia Karadoulama, Dawid Mąkosa, Michal Lubas, Torben Heick Jensen

Correspondence

thj@mbg.au.dk

In Brief

Silla et al. report that the RNA exosome adaptor protein ZFC3H1 acts as a nuclear RNA retention factor. In the absence of ZFC3H1, exosome targets are exported to the cytoplasm in a AlyREF-dependent manner. The discovery establishes ZFC3H1 as a central factor in the retention and degradation of polyadenylated RNA.

Data and Software Availability

GSE108197

Silla et al., 2018, Cell Reports *23*, 2199–2210
May 15, 2018
https://doi.org/10.1016/j.celrep.2018.04.061

CellPress

Cell Reports
Article

The RNA Exosome Adaptor ZFC3H1 Functionally Competes with Nuclear Export Activity to Retain Target Transcripts

Toomas Silla,[1] Evdoxia Karadoulama,[1,2] Dawid Mąkosa,[1] Michal Lubas,[1,3] and Torben Heick Jensen[1,4,*]
[1]Department of Molecular Biology and Genetics, Aarhus University, C.F. Møllers Allé 3, Building 1130, 8000 Aarhus C, Denmark
[2]The Bioinformatics Centre, Department of Biology and Biotech Research and Innovation Centre, University of Copenhagen, Ole Maaloesvej 5, 2200 Copenhagen, Denmark
[3]Present address: Biotech Research and Innovation Centre, University of Copenhagen, 2200 Copenhagen N, Denmark
[4]Lead Contact
*Correspondence: thj@mbg.au.dk
https://doi.org/10.1016/j.celrep.2018.04.061

SUMMARY

Mammalian genomes are promiscuously transcribed, yielding protein-coding and non-coding products. Many transcripts are short lived due to their nuclear degradation by the ribonucleolytic RNA exosome. Here, we show that abolished nuclear exosome function causes the formation of distinct nuclear foci, containing polyadenylated (pA$^+$) RNA secluded from nucleocytoplasmic export. We asked whether exosome co-factors could serve such nuclear retention. Co-localization studies revealed the enrichment of pA$^+$ RNA foci with "pA-tail exosome targeting (PAXT) connection" components MTR4, ZFC3H1, and PABPN1 but no overlap with known nuclear structures such as Cajal bodies, speckles, paraspeckles, or nucleoli. Interestingly, ZFC3H1 is required for foci formation, and in its absence, selected pA$^+$ RNAs, including coding and non-coding transcripts, are exported to the cytoplasm in a process dependent on the mRNA export factor AlyREF. Our results establish ZFC3H1 as a central nuclear pA$^+$ RNA retention factor, counteracting nuclear export activity.

INTRODUCTION

Mammalian genomes transcribe a large repertoire of RNAs, which experience markedly different fates. This includes (1) their transport to nuclear compartments for functioning or for the passage of processing steps, (2) their nuclear export, or (3) their rapid nuclear decay. Nuclear export of RNA polymerase II (Pol II) transcripts is tightly coupled to their 5′ m^{7}G cap formation, splicing, and 3′polyadenylation. mRNAs serve as the prototypic example of such flow from transcription to cytoplasmic translation (recently reviewed in Williams et al. [2018]). Some long non-coding RNAs (lncRNAs) are processed similarly to mRNAs (Kung et al., 2013). However, many are also rapidly degraded in the nucleus without being extensively spliced and where the polyadenylation status is unclear. This includes divergently transcribed enhancer RNAs (eRNAs) (Andersson et al., 2014; Djebali et al., 2012) and so-called promoter upstream transcripts (PROMPTs)/upstream antisense (UA) RNAs (Preker et al., 2008; Seila et al., 2008). Despite of all these RNAs being Pol II products and displaying structural similarity, at least at their 5′ ends, lncRNAs appear to be more concentrated in the nucleus and to be more labile than their mRNA counterparts (Djebali et al., 2012; Preker et al., 2011). Thus, mechanisms must exist that "sort" Pol II-derived transcripts for their differential nuclear and cytoplasmic destinies.

RNA-processing steps largely take place co-transcriptionally but, nevertheless, may occur in specialized nuclear areas; nucleoli, speckles, Cajal bodies (CBs), and paraspeckles (Sleeman and Trinkle-Mulcahy, 2014). Commonly, such nuclear bodies are self-organizing and dynamic structures that harbor specific substrates and enzymes and where enzymatic and RNP assembly reactions might be accelerated. With relevance for Pol II-derived transcripts, CBs contain a variety of proteins and RNAs that are involved in the assembly and modification of small nuclear ribonucleoproteins/small nucleolar ribonucleoproteins (snRNPs/snoRNPs) (Nizami et al., 2010). After their maturation, snRNPs travel to speckles, which may host multiple steps of gene expression, such as Pol II transcription, pre-mRNA splicing, and other mRNP maturation events in preparation for nuclear export (recently reviewed in Galganski et al. [2017]). Paraspeckles, which are proximal to speckles, form around the NEAT1 lncRNA and are enriched for specific RNA-binding proteins (Naganuma and Hirose, 2013). Their function is still debated, but recent reports suggest that paraspeckle components are involved in gene expression regulation (West et al., 2014), immune responses (Morchikh et al., 2017), and microRNA processing (Jiang et al., 2017). The so-called "Mmi foci" are another group of nuclear bodies of interest. These foci, which, so far, appear specific for *S. pombe* cells, are enriched with proteins involved in nuclear RNA decay, and it has been proposed that Mmi1 foci are degradation sites for meiosis-specific transcripts during vegetative growth (Harigaya et al., 2006; Sugiyama and Sugioka-Sugiyama, 2011; Yamanaka et al., 2010; Yamashita et al., 2013). Where and whether similar degradation centers exist in mammalian cell nuclei are still open questions.

The best described and major, nuclear RNA decay machinery is the highly conserved 3′–5′ exo- and endo-nucleolytic RNA exosome (Kilchert et al., 2016). This multi-subunit complex is

present in the cytoplasm as well as in the nucleus, where it deals with most RNA biotypes. The human nucleoplasmic exosome is composed of a nine-subunit catalytically inactive core (known as EXO9) that achieves its activity via interaction with the exonuclease RRP6 and the exo- and endonuclease RRP44 (also known as DIS3) (Kilchert et al., 2016; Zinder and Lima, 2017). RNA exosome function critically depends on the RNA helicase MTR4 (also called SKIV2L2) to unwind exosome substrates, hereby facilitating their entry into the central channel of the exosome (Johnson and Jackson, 2013; Schneider and Tollervey, 2014). Critically, MTR4 also serves as a platform for recruiting adaptor proteins, providing RNA substrate binding and specificity (Meola and Jensen, 2017). In human nuclei, three such MTR4-containing adaptor complexes have been described. The human hTRAMP complex, harboring the non-canonical poly(A) (pA) polymerase TRF4-2 and the zinc-knuckle protein ZCCHC7, is exclusively nucleolar and predominantly involved in rRNA processing and decay (Lubas et al., 2011). Two other MTR4-containing complexes, the nuclear-exosome-targeting (NEXT) complex (Lubas et al., 2011) and the pA-tail-exosome-targeting (PAXT) connection (Meola et al., 2016), both have nucleoplasmic localizations. In the NEXT complex, MTR4 interacts with the zinc-knuckle protein ZCCHC8 and the promiscuous RNA-binding protein RBM7 (Giacometti et al., 2017; Lubas et al., 2011), which facilitate exosome decay of, e.g., PROMPTs, eRNAs, and 3′ extended products of snRNAs and snoRNAs (Andersen et al., 2013; Hrossova et al., 2015; Lubas et al., 2011, 2015; Meola et al., 2016). In case of the PAXT connection, another zinc-finger protein, ZFC3H1, mediates an interaction of MTR4 with the nuclear pA-binding protein PABPN1, in turn, targeting polyadenylated RNAs to the exosome (Meola et al., 2016). Thus, although substrate overlap exists, PAXT generally promotes the degradation of longer and more extensively polyadenylated RNAs, whereas NEXT targets shorter and more immature transcripts (Meola et al., 2016).

Adding an m^7G cap to the 5′ end of the ~20-nt-long nascent transcripts is considered to be a hallmark of successful Pol II transcription initiation (Ramanathan et al., 2016). Shortly after its capping, the nascent RNA is bound by the cap-binding proteins, CBP20 and CBP80, forming the cap-binding complex (CBC) (Izaurralde et al., 1994; Moteki and Price, 2002; Visa et al., 1996). Interestingly, both PAXT and NEXT components can connect to the CBC via bridging proteins ZC3H18 and ARS2 (Andersen et al., 2013; Hallais et al., 2013; Meola et al., 2016). ARS2 directly interacts with the CBC to form the CBC-ARS2 (CBCA) complex, which behaves as a general suppressor of pervasive transcription (Andersen et al., 2013; Iasillo et al., 2017), whereas ZC3H18 connects the CBCA with NEXT for immediate substrate degradation (Andersen et al., 2013). The CBC also plays a role in intra-nuclear transcript transport as well as nuclear RNA export (Boulon et al., 2004; Cheng et al., 2006). For small nuclear RNAs (snRNAs), the transport adaptor protein PHAX binds the CBCA complex to promote the nuclear export of capped precursor transcripts (Boulon et al., 2004; Ohno et al., 2000). Sorting these RNAs between nuclear decay and export is suggested to occur in a competition between ZC3H18/NEXT and PHAX for CBCA binding (Giacometti et al., 2017). For most polyadenylated RNAs, like mRNAs, nuclear export is promoted in a process involving the AlyREF protein, which interacts with CBP80 to recruit the transcription/export (TREX) machinery to mRNA 5′ ends (Masuda et al., 2005). How the interplay between RNA export and degradation factors leads to a final decision of RNA fate remains unresolved.

In this study, we demonstrate that abolishing RNA exosome activity, by depleting its core component RRP40 (also known as EXOSC3), leads to the sequestration of pA^+ RNAs in distinct nuclear foci (coined pA^+ RNA foci). We find that MTR4 and PAXT, but not NEXT, components, are enriched in pA^+ RNA foci, and we identify ZFC3H1 as a critical factor for foci formation. We surmise that ZFC3H1 is a nuclear retention factor for its bound pA^+ RNA, and we demonstrate that it functionally competes with the pA^+ RNA export factor AlyREF.

RESULTS

RNA Exosome Depletion Leads to the Accumulation of pA^+ RNA in Distinct Nuclear Foci

While both NEXT and PAXT pathways act on nuclear transcripts, it is unclear where targeting occurs and which mechanism(s) might ensure the nuclear retention of substrates prior to their decay. To begin to shed light on these questions, we first used small interfering RNAs (siRNAs) to deplete the exosome core protein RRP40 in HeLa cells (Figure 1A) and visualized pA^+ RNA with a fluorescently labeled oligo(dT) probe. By removing an exosome core component, all exosome targets are impacted, regardless of the utilized adaptor complex. Moreover, it has been demonstrated that, in the absence of the exosome, at least some of its targets undergo poly- (Preker et al., 2011) or hyper-adenylation (Bresson and Conrad, 2013). RRP40 depletion triggered the appearance of distinct nuclear pA^+ RNA foci only vaguely visible in control cells (Figure 1B). To ascertain whether these foci were related to known nuclear bodies, we performed co-localization analysis using SC35, Coilin, and PSF protein as markers for speckles, paraspeckles, and Cajal bodies, respectively. These analyses did not disclose any exclusive overlap with pA^+ RNA foci in exosome-depleted conditions (Figures 1C, S1A, and S1B), although we noted an occasional co-localization with SC35 (Figure 1C, right line scan) and PSF (Figure S1B, right line scan) and a frequent close proximity to Coilin (Figure S1A, right line scan). In addition, we did not observe pA^+ RNA localization in nucleoli, constituting condensed regions poorly stained with DAPI. We conclude that at least some exosome targets are nuclear retained and largely aggregate outside of known nuclear compartments when prevented from getting degraded.

PAXT Components Are Selectively Enriched in pA^+ RNA Foci

Besides RNA, nuclear bodies harbor high concentrations of protein, potentially involved in their biological function. Therefore, we asked whether pA^+ RNA foci contain proteins involved in nuclear RNA decay. Immunolocalization (IL) analysis of MTR4, involved in both NEXT and PAXT pathways, displayed a clear punctate accumulation in RRP40-depleted cells, whereas the protein showed an even nuclear distribution in control cells (Figure 2A). Importantly, dual MTR4 IL/pA^+ RNA fluorescence in situ

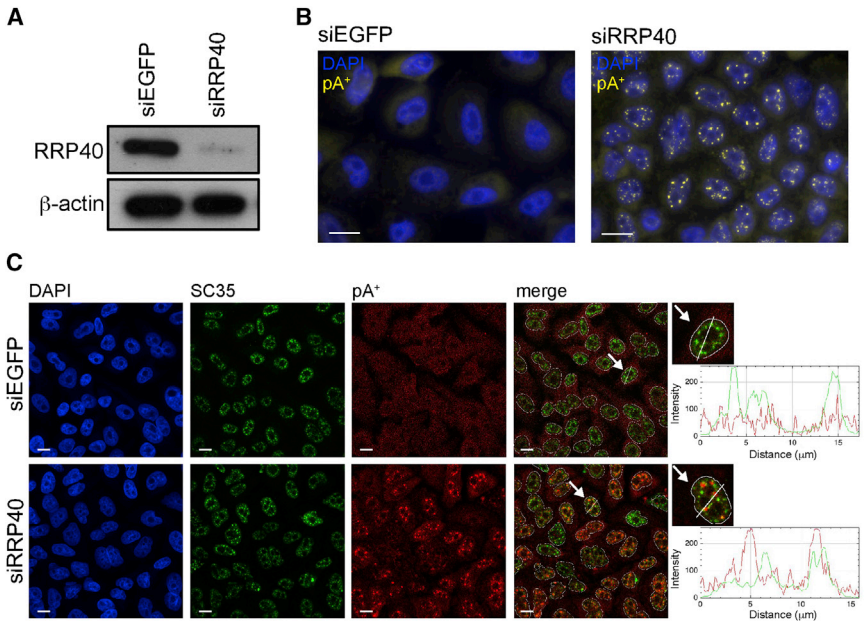


Figure 1. RNA Exosome Depletion Leads to the Formation of Distinct Nuclear pA+ RNA Foci

(A) Western blotting analysis of RRP40 levels in HeLa cells treated with control siRNA (siEGFP) or siRNA targeting RRP40 (siRRP40). β-actin was used as a loading control.

(B) RNA-FISH analysis of pA+ RNA in control (siEGFP) or RRP40-depleted (siRRP40) HeLa cells. pA+ RNA, detected with an oligo(dT)-LNA probe (Thomsen et al., 2005), is shown in yellow, and nuclear DAPI stain is shown in blue.

(C) Co-localization analysis of pA+ RNA foci with the speckle marker SC35 using cells from (B). Confocal images of SC35 immunofluorescence (IF) signal in green, pA+ RNA-FISH signal in red, and DAPI stain in blue. SC35 IF and RNA-FISH signals were merged (merge). Arrows point to cells that were used for line scan analyses, and zoom-ins of the relevant cells are shown at the right of the corresponding panel. Line scan profiles represent IF and RNA-FISH signal intensities along the drawn line. Dashed lines represent outlines of nuclei. pA+ RNA was detected as in (B).

Scale bars, 10 μm.

hybridization (FISH) analysis demonstrated that MTR4 punctae coincide with pA+ RNA foci (Figure 2A, right bottom line scan).

To inquire about PAXT and/or NEXT enrichment in the pA+ RNA foci, we repeated such co-localization analysis using antibodies toward the PAXT components PABPN1 and ZFC3H1, as well as the NEXT components RBM7 and ZCCHC8. This revealed a strong overlap between PABPN1 and ZFC3H1 with the pA+ RNA signal upon exosome inactivation (Figures 2B and 2C, right bottom line scans), whereas neither of the NEXT components ZCCHC8 and RBM7 exhibited any noticeable localization change between RRP40-depleted and control cells (Figures S2A and S2B). Interestingly, we could also detect an enrichment of the ARS2 protein in pA+ RNA foci (Figure S2C). As ARS2 is part of the CBCA complex that binds capped RNA, this suggests that accumulating RNAs have a proper cap structure. Thus, PAXT components MTR4, ZFC3H1, and PABPN1 accumulate on nuclear exosome targets, of which at least some are capped and polyadenylated, when turnover is prevented.

Diverse RNAs Accumulate in pA+ RNA Foci upon Exosome Inactivation

Having established an enrichment of PAXT proteins in pA+ RNA foci, we next addressed whether known RNA substrates of the PAXT pathway were also present. Hence, we performed specific RNA-FISH analysis of the highly expressed and spliced SNHG19 RNA previously established as a PAXT pathway target (Meola et al., 2016) (Figure S3A). FISH analysis using probes spanning the SNHG19 RNA exon-exon junction demonstrated its co-accumulation with polyadenylated RNA in pA+ RNA foci (Figure 3A). To determine any pathway specificity of such accumulation, we next localized NEXT pathway targets. These are generally short and lowly expressed, making RNA-FISH challenging; therefore, we chose the highly expressed DNAJB4 PROMPT (Figure S3B) and the 3′ extended form of U1 snRNA (RNVU1-14) (Figure S3C) for our analysis. Both these transcripts accumulated in pA+ RNA foci upon RNA exosome depletion (Figures 3B and 3C). Finally, we analyzed RNA expressed from an artificial PROMPT (proPOGZ) locus, stably integrated in the HeLa genome between a CMV promoter and a BGH pA signal (Ntini et al., 2013). qRT-PCR analysis showed this transcript to be stabilized upon RNA exosome depletion, although neither ZCCHC8 (NEXT) nor ZFC3H1 (PAXT) depletion affected its levels (Figure S3D). Yet, RNA-FISH analysis still revealed a clear accumulation of this transcript in pA+ RNA foci (Figure S3E). Thus, even though the nuclear pA+ foci specifically concentrate PAXT components, they contain both PAXT and non-PAXT RNA substrates. It is plausible that non-PAXT targets become polyadenylated in exosome-depletion conditions (Preker et al., 2011) and that longer pA tails trigger PAXT loading and pA+ RNA foci localization (see also Discussion).

The ZFC3H1 Protein Is Required for Formation of pA+ RNA Foci

Given the massive condensation of nuclear RNA exosome substrates into pA+ RNA foci, we next wondered whether exosome adapters NEXT or PAXT might be involved in foci formation. To examine this question, we depleted HeLa cells of RRP40, MTR4, ZFC3H1, or ZCCHC8 (Figure S4A) and followed pA+ RNA foci and SNHG19 RNA localization by RNA-FISH, as described earlier. As previously reported (Lubas et al., 2011; Ogami et al., 2017), MTR4 depletion also reduced ZCCHC8, ZFC3H1, and, to some extent, RRP40 protein levels (Figure S4A). Still, neither MTR4 nor ZFC3H1 and ZCCHC8 depletions resulted in pA+ or SNHG19 RNA accumulation in nuclear foci (Figures 4A and S4B). At the same time, parallel qRT-PCR analysis showed the robust accumulation of SNHG19 RNA upon both MTR4 and ZFC3H1 depletions (Figure 4B; see also Meola et al. [2016]).

A possible reconciliation of these data could be that the MTR4 and/or ZFC3H1 proteins themselves are, apart for their

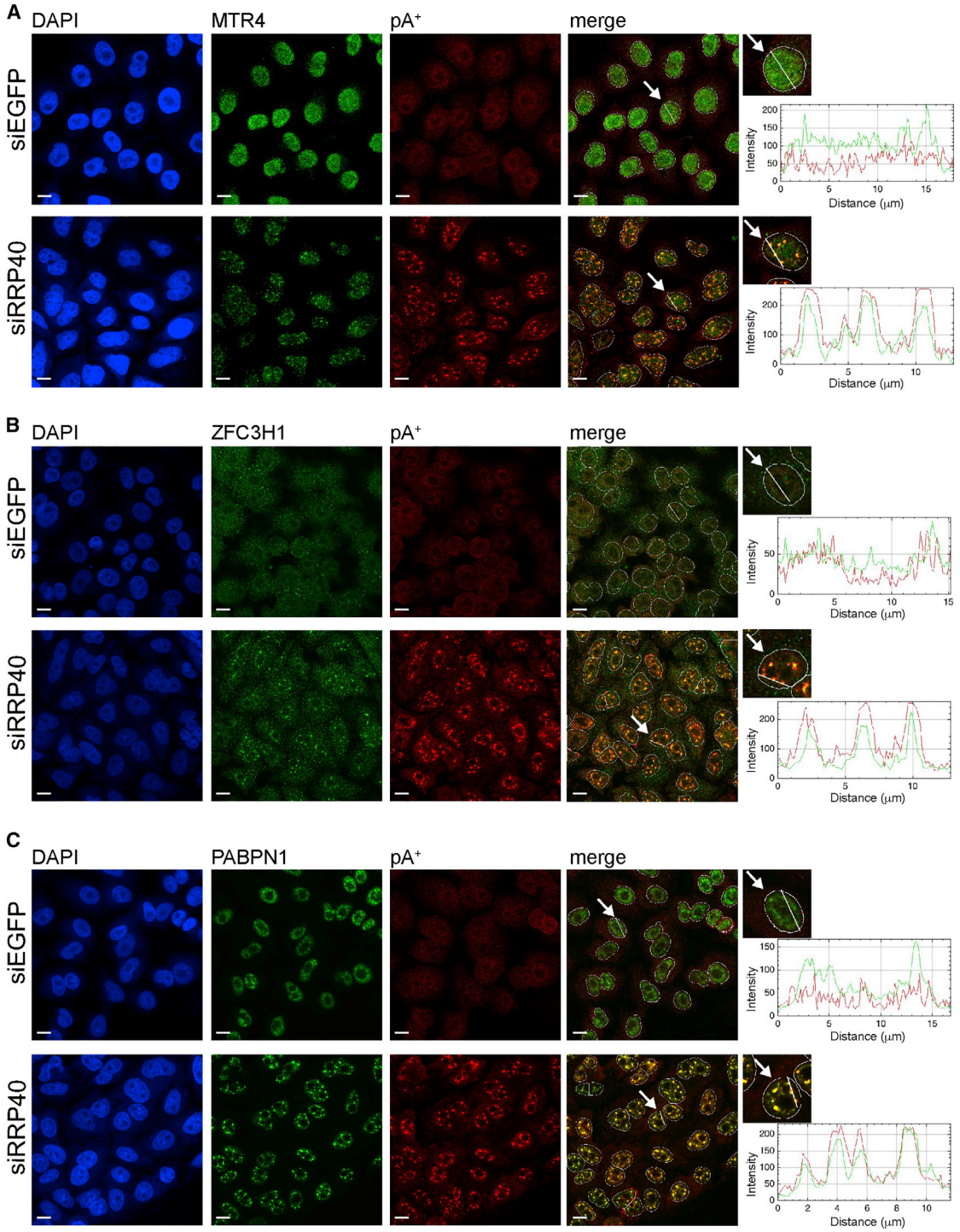


Figure 2. PAXT Complex Components Localize to pA+ RNA Foci

(A–C) Co-localization analysis of PAXT components MTR4 (A), ZFC3H1 (B), and PABPN1 (C) with pA+ RNA in control (siEGFP) and RRP40-depleted (siRRP40) HeLa cells as in Figure 1C. Scale bars, 10 μm.

roles as exosome co-factors, required for the creation of pA+ RNA foci. To test this notion, we depleted cells for RRP40 (to form pA+ RNA foci) and combined this with depletion of MTR4, ZFC3H1, or ZCCHC8, yielding largely comparable levels of SNHG19 RNA as upon single depletion of RRP40 (Figure 4B). Parallel RNA-FISH analysis showed that cells subjected to

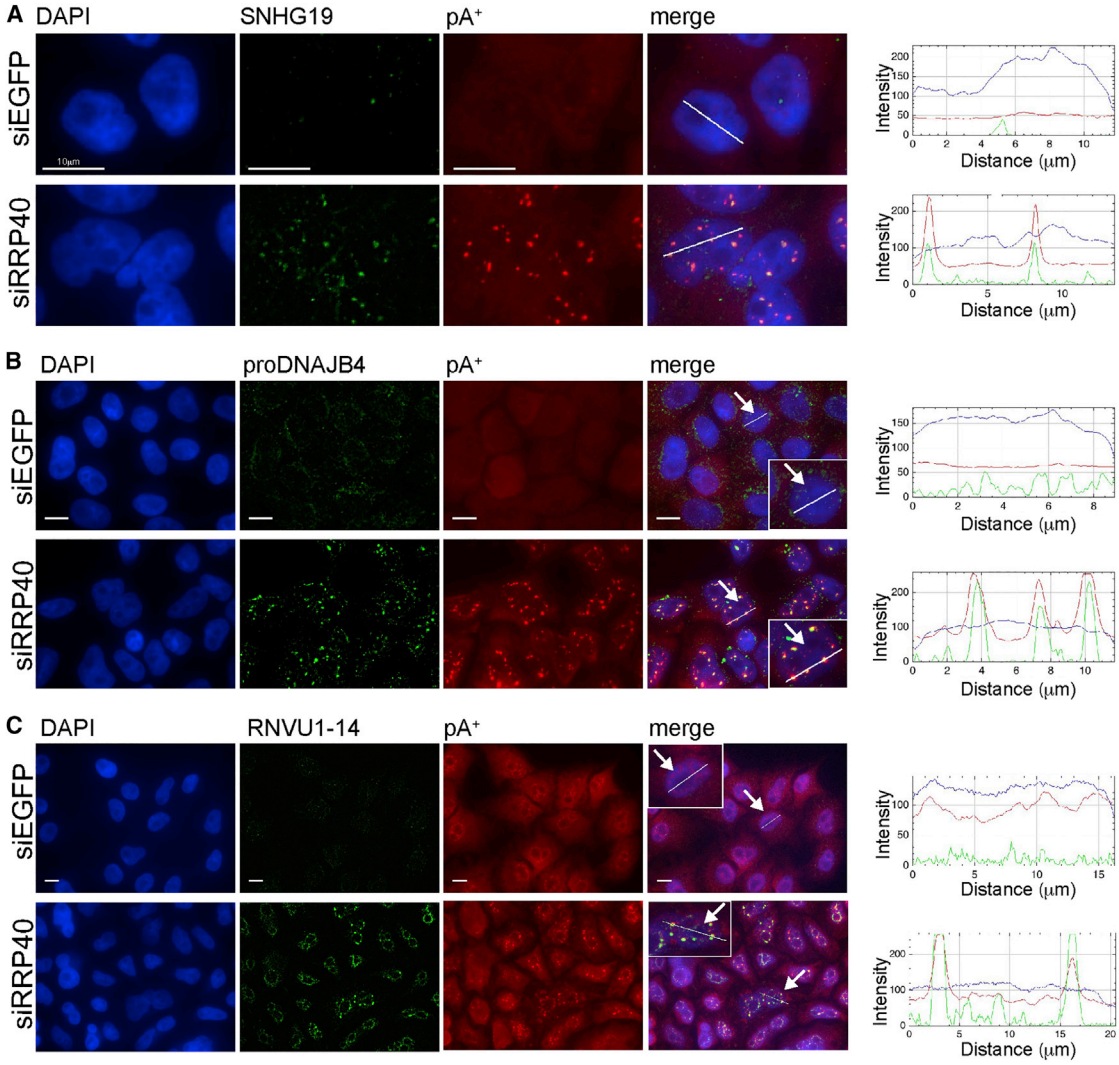


Figure 3. Diverse RNAs Accumulate in pA$^+$ RNA Foci

(A–C) RNA-FISH co-localization analyses of pA$^+$ RNA with SNHG19 (A), proDNAJB4 (B), and RNVU1-14 3′ extension (C) in control (siEGFP) or RRP40-depleted (siRRP40) HeLa cells. pA$^+$ RNA was detected with an oligo(dT) 50-mer probe. Line scans are indicated as in Figure 1C, including DAPI channel. Scale bars, 10 μm.

RRP40/MTR4 and RRP40/ZFC3H1 co-depletions were devoid of any discernable pA$^+$ and SNHG19 RNA accumulation (Figure 4A). In contrast, cells subjected to RRP40/ZCCHC8 co-depletion displayed similar focal patterns of pA$^+$ and SNHG19 RNA, as observed for cells depleted only for RRP40. Given the presence of MTR4 protein in RRP40/ZFC3H1-depleted cells and the co-depletion of MTR4 and ZFC3H1 in RRP40/MTR4-depleted cells (Figure S4A), these results, therefore, implicate the ZFC3H1 protein as an important contributor to pA$^+$ RNA foci formation in exosome-depleted cells. If so, exogenous ZFC3H1 expression in RRP40/ZFC3H1 co-depleted cells should re-install pA$^+$ foci formation. To test this, we used ZFC3H1 5′ and 3′ UTR-specific siRNAs, which both efficiently deplete endogenous ZFC3H1 protein but do not target a plasmid-expressed FLAG-tagged ZFC3H1 version (Figure S4C). Indeed, parallel IL/pA$^+$ RNA-FISH analysis of the plasmid-transfected cells demonstrated that pA$^+$ foci formation was rescued by exogenous FLAG-tagged ZFC3H1 expression (Figure S4D). We conclude that ZFC3H1 is required for pA$^+$ RNA foci formation.

What then is the fate of exosome substrates, e.g., SNHG19 RNA, in ZFC3H1-depleted cells, where the transcript is stabilized but does not accumulate in nuclear foci? If ZFC3H1 activity generally retains exosome substrates in cell nuclei, SNHG19 transcripts escaping such retention would be likely to enter the mRNA export pathway. To explore this possibility, we performed subcellular fractionation of HeLa cells from Figure 4A and analyzed the abundance of SNHG19 RNA in nuclear and cytoplasmic fractions. qRT-PCR analysis showed a clear accumulation of SNHG19 transcripts in the cytoplasmic fractions of cells subjected to factor depletion combinations; MTR4, ZFC3H1, RRP40/MTR4, and RRP40/ZFC3H1, which all yielded

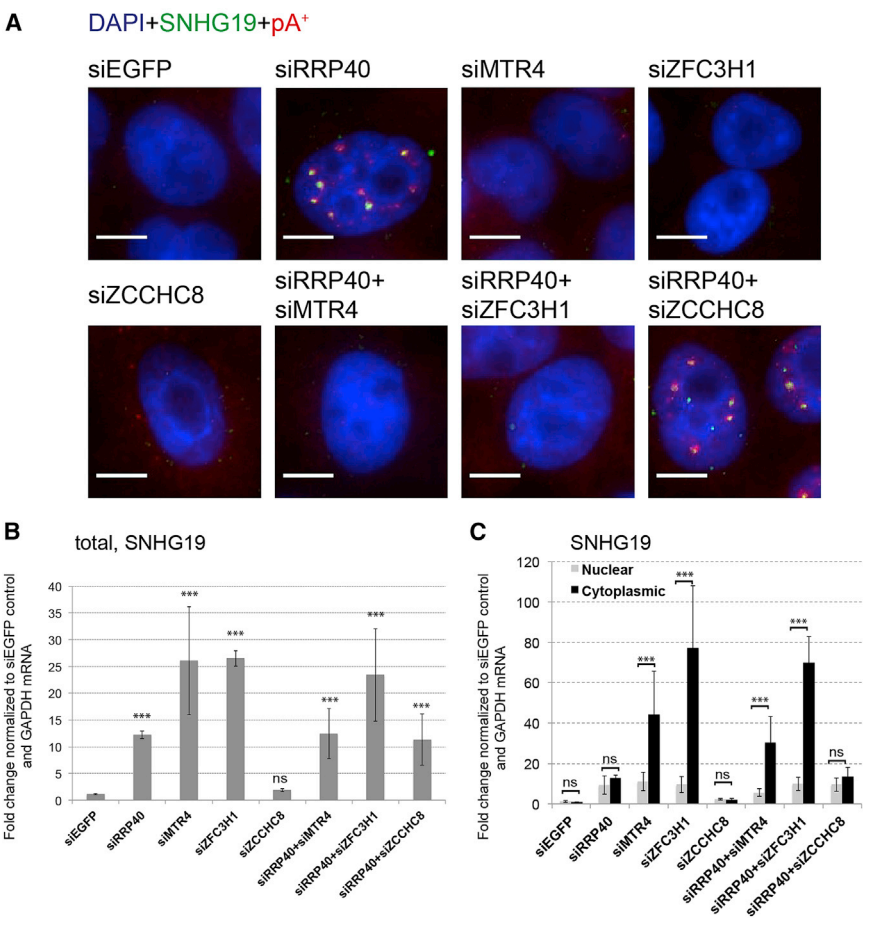


Figure 4. The ZFC3H1 Protein Is Required for RNA Foci Formation

(A) Dual SNHG19 and pA$^+$ RNA-FISH analysis in the indicated factor-depleted HeLa cells. Merged images of DAPI (blue), SNHG19 (green), and pA$^+$ RNA (red) are shown. Images from separate channels are shown in Figure S4B. pA$^+$ RNA was detected with an oligo(dT) 50-mer probe.

(B and C) qRT-PCR analysis of SNHG19 RNA using total (B) or cytoplasmic and nuclear (C) fractions harvested from the indicated factor-depleted HeLa cells. In (C), 1 μg RNA from the respective fractions was used for reverse transcriptase reactions. Data are displayed as mean values, with error bars denoting SD (n = 4 biological replicates, except for siZCCHC8: n = 3). ***p < 0.01, Student's t test; ns, not significant. Scale bars, 10 μm.

See also Figure S4B.

low ZFC3H1 levels (Figures 4C and S4A). Consistently, ZCCHC8 depletion alone, or in combination with RRP40 depletion, did not result in elevated cytoplasmic SNHG19 RNA levels (Figure 4C). These results all point toward ZFC3H1 being responsible for retaining RNA exosome targets in the cell nucleus.

The mRNA Export Factor AlyREF Functionally Competes with ZFC3H1

Our results so far suggest that ZFC3H1 might compete with RNA export activity to prevent nuclear RNA exosome substrates from entering the cytoplasm. IL analysis of mRNA export factors AlyREF (Figures S5A and S5B) and NXF1 (data not shown) did not reveal any accumulation in pA$^+$ RNA foci, suggesting mutual exclusive binding of export and decay factors. Therefore, the observed export of exosome substrates upon ZFC3H1 depletion may be caused by recruitment of the export machinery to substrates normally blocked by decay co-factors. To test this prediction, we depleted AlyREF alone or in ZFC3H1- and RRP40/ZFC3H1-depletion conditions (Figure S5C). As expected, depletion of AlyREF alone resulted in nuclear accumulation of pA$^+$ RNA (Figure 5A) but did not lead to nuclear accumulation of SNHG19 RNA (Figure 5B), which is most likely degraded by the exosome in this condition. When co-depleted with ZFC3H1 or RRP40/ZFC3H1, AlyREF depletion resulted in re-established nuclear accumulation of SNGH19 and pA$^+$ RNA (Figures 5A and 5B). However, as opposed to the distinct foci of RRP40-depleted cells, pA$^+$ and SNHG19 RNA now appeared more diffusely in the nucleoplasm (Figures 5A and 5B; compare panel "siRRP40" to panels "siZFC3H1/siAlyREF" and "siRRP40/siZFC3H1/siAlyREF"). These results reiterate a role of ZFC3H1 in pA$^+$ RNA aggregation and suggest that the protein might "shield" RNA exosome targets from the mRNA export machinery. In the absence of ZFC3H1, however, even exosome targets with no normal cytoplasmic business, like SNHG19 RNA, will get exported from the nucleus.

Full-Length mRNAs Accumulate upon Exosome Depletion

To further examine which transcripts get turned over by the nuclear exosome and might accumulate in pA$^+$ RNA foci in its absence, we performed high-throughput RNA-sequencing (RNA-seq) analysis of pA$^+$ RNA from nuclear fractions of cells subjected to RRP40 depletion ("siRRP40") or treated with control EGFP siRNAs ("siEGFP") (Figure S6A). Both types of libraries were produced either in duplicates (EGFPkd) or triplicates (RRP40kd), which generally displayed good reproducibility between replicates (Figure S6B). Nuclear exosome targets were revealed by calculating fold changes between the RRP40kd and the control sample, using DESeq2, and using all GENCODE-annotated RNAs expressed in the samples as well as PROMPTs and eRNAs as previously described (Meola et al., 2016). As expected PROMPTs and eRNAs were generally upregulated in the RRP40-depletion condition (Figure 6A), consistent with the notion that a large share of these transcripts are adenylated (Preker et al., 2011). More surprisingly, we noticed that RNA stemming from a significant number of annotated protein-coding genes were either up- or downregulated upon RRP40 depletion (Figure S6C), which was also observed

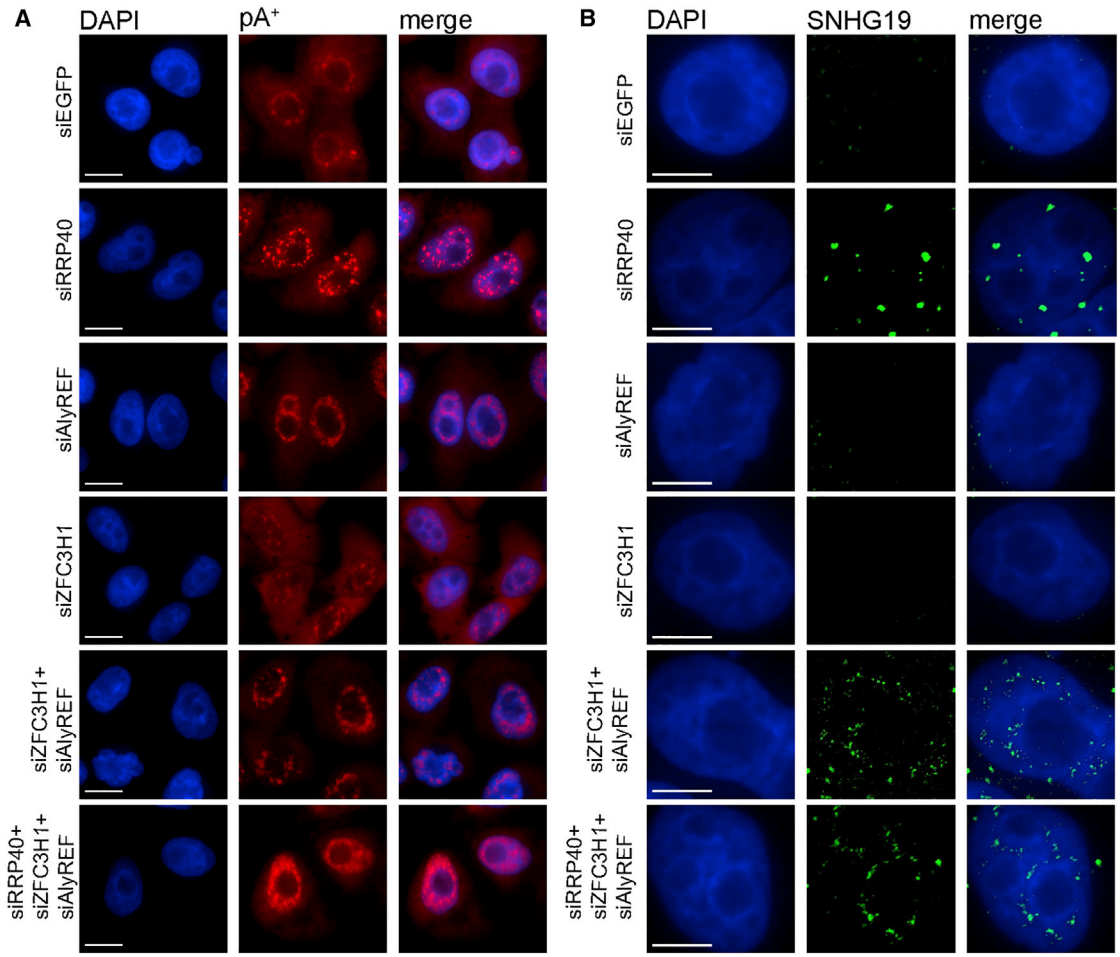


Figure 5. ZFC3H1 Counteracts AlyREF Function

(A and B) RNA-FISH analysis of pA^+ RNA (A) and SNHG19 RNA (B) in factor-depleted (indicated to the left side of each image panel) HeLa cells. Image display is as shown in Figure 3B. pA^+ RNA was detected with an oligo(dT) 50-mer probe. Scale bars, 10 μm.

in previously published RRP40kd libraries of total RNA preparations (Meola et al., 2016) (Figure S6D). However, aberrant products, e.g., those constituting premature transcription termination events, as well as intragenic eRNAs are stabilized upon RRP40 depletion (Iasillo et al., 2017). Therefore, we intersected the nuclear RNA-seq analysis with our previously published total RRP40kd RNA-seq data (Meola et al., 2016) and also interrogated the extent to which the upregulated signals constituted full-length and/or spliced mRNA. To this end, we selected 1,641 protein-coding genes that were significantly upregulated (log_2[RRP40kd/EGFPkd] > 0, and adjusted p value [padj] < 0.05) in the nuclear samples and were upregulated/unaffected (log_2[RRP40kd/EGFPkd] > −0.5) in the total RNA datasets (Figure 6B). Log_2 fold changes (RRP40kd/EGFPkd) were then calculated for each exon and intron, and genes were selected that had their exons upregulated in the nuclear RNA samples, resulting in a set of 673. On average, introns were less upregulated than exons, indicating that mature spliced RNAs were being stabilized by exosome depletion (Figure 6C). Moreover, upregulation occurred across all exons, strongly suggesting that full-length mRNAs were affected. We note that our analysis cannot discriminate which of these are direct targets of the exosome and which accumulate due to indirect effects.

To characterize exosome-sensitive mRNAs, we used a published classification system for human coding and non-coding genes (Mukherjee et al., 2017). In relation to a background set (Figure S6E), our set of 1,641 upregulated protein-coding genes was significantly enriched for cluster c3 (p = 6.12e−10) from Mukherjee et al. (2017) (Figure 6D). A significant enrichment (p = 3.207e−07) was obtained for the same cluster when analyzing only those 366 genes giving rise to mRNAs upregulated (log_2[RRP40kd/EGFPkd] > 0, and padj < 0.05) in both total and nuclear samples (Figures S6F and S6G). Interestingly, cluster c3 is enriched for transcription-factor-encoding mRNAs, that are more nuclear than the average, are well synthesized, and exhibit high degradation rates (Mukherjee et al., 2017).

Finally, we asked whether exosome-sensitive mRNAs might also localize to pA^+ RNA foci upon RRP40 depletion. The ACTB and NFKB2 mRNAs were selected for analysis, as both transcripts are relatively abundant, and their full-length mature

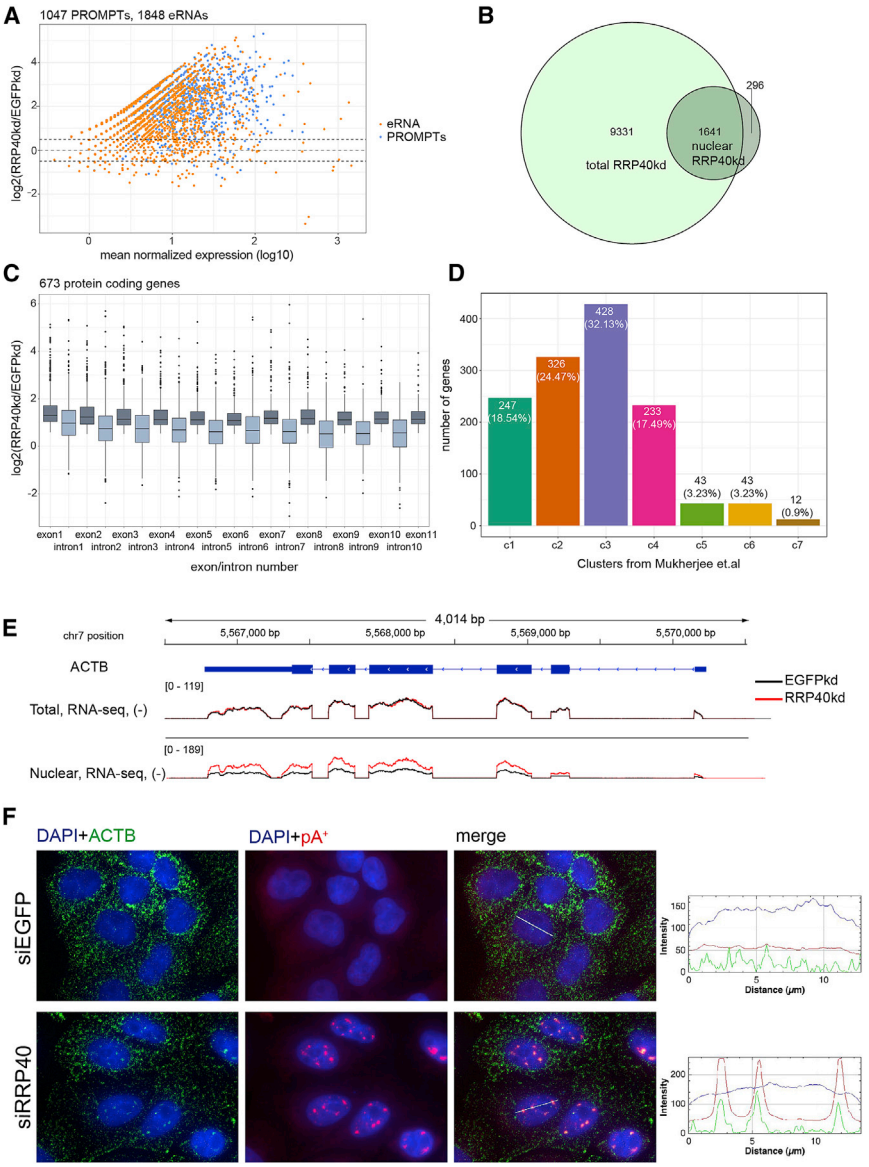


Figure 6. Full-Length mRNAs Can Be Upregulated upon Exosome Depletion

(A) Log_2 fold changes (y axis) between PROMPTs and eRNAs in nuclear RRP40kd versus EGFPkd libraries plotted against their mean normalized expression (log_{10} scale, x axis).

(B) Venn diagram representing the overlap between significantly upregulated protein-coding genes (log_2[RRP40kd/EGFPkd] > 0, padj < 0.05) in the nuclear RNA libraries and the upregulated and unaffected protein-coding genes (log_2[RRP40kd/EGFPkd] > −0.5) in the total RNA libraries.

(C) Boxplots of log_2 fold changes (RRP40kd/EGFPkd) of each individual exon and intron of the selected protein coding genes in ascending order up to the 11th exon.

(D) Bar plots representing the classification of 1,641 protein-coding genes—overlap from (B)—to previously described clusters (Mukherjee et al., 2017). Note the enrichment of genes belonging to cluster c3 (Pearson's chi-square test, p = 6.12e−10; calculated against all expressed genes in nuclear samples represented in Figure S6E).

(E) Genome browser view of the ACTB locus, showing tracks of normalized total and nuclear RNA-seq data from control (EGFPkd, black line) and RRP40-depleted (red line) HeLa cells. All tracks represent an average of three biological replicates, except the nuclear EGFPkd sample, which represents average of two biological replicates.

(F) Dual ACTB and pA^+ RNA-FISH analysis in the indicated factor-depleted HeLa cells. Merged images of DAPI (blue), ACTB (green), and pA^+ RNA (red) are shown. Image display is as in Figure 1C.

See also Figure S6E.

RNA species displayed elevated nuclear expression upon RRP40 depletion (Figures 6E, S6H, and S7A). As expected, RNA-FISH experiments revealed a predominant cytoplasmic localization for both RNAs, which did not differ between control and RRP40-depleted cells (Figures 6F and S7B). However, some ACTB mRNA could be detected in nuclear pA^+ RNA foci upon RRP40 depletion (Figure 6F), while a noticeable accumulation of NFKB2 RNA was not observed (Figure S7B). Collectively, these results suggest that a fraction of spliced, full-length mRNA is subject to nuclear exosomal decay.

DISCUSSION

The vast majority of the human genome is transcribed, and commonly from non-genic regions, giving rise to a broad variety of ncRNA. These are often transcribed by Pol II and, therefore, pass some processing steps characteristic of mRNA. Yet, unlike most mRNA, a large fraction of ncRNA is rapidly turned over by the nuclear RNA exosome. How such molecular decision is enabled, and with what means nuclear export of these ncRNA is prevented, is not well understood. Here, we demonstrate that precluding the degradation of RNA exosome targets, including some mRNA, does not automatically lead to their nuclear export. Instead, they form distinct nuclear pA^+ RNA foci, requiring retention that counteracts export. We unveil one aspect of such mechanism(s) by demonstrating that the PAXT component ZFC3H1 is required for establishing the pA^+ RNA foci phenotype. Moreover, we suggest that this occurs in functional competition with AlyREF-mediated pA^+ RNA nuclear export.

Cell biological data presented here echo results from mitotically growing *S. pombe* cells, where unwanted meiosis-specific transcripts are cleared by the nuclear exosome (Chen et al., 2011; Harigaya et al., 2006; Hiriart et al., 2012; Yamashita et al., 2012; Zofall et al., 2012). These transcripts are specifically recognized by the Mmi1 protein (Harigaya et al., 2006) in the context of MTREC, a multisubunit complex that also contains Red1 and Pab2, which are *S. pombe* homologs of ZFC3H1

and PAPBN1, respectively (Marayati et al., 2016; Zhou et al., 2015). Both these proteins are required for the decay of meiotic transcripts, in a process that also requires the polyadenylation activity of the canonical pA polymerase Pla1 (St-André et al., 2010; Sugiyama and Sugioka-Sugiyama, 2011; Yamanaka et al., 2010). Interestingly, Mmi1, Red1, Pab2, polyadenylation factors, and the nuclear exosome subunit Rrp6 accumulate in distinct nuclear dots, suggested to be transcript degradation sites (Harigaya et al., 2006; Sugiyama and Sugioka-Sugiyama, 2011; Yamanaka et al., 2010). Given the obvious similarity to our results, it is possible that such RNA decay centers also exist in mammalian cells. However, unlike in *S. pombe*, where "Mmi1 foci" are also visible in wild-type cells, we do not measure discernable condensates of exosome adaptor proteins in normally growing HeLa cells. Although we cannot fully exclude that RNA foci are pre-existing in normal cells as condensates of non-adenylated RNA, which becomes adenylated after RNA exosome depletion, the fact that neither adaptor proteins nor targets of the exosome accumulate in foci in the control condition is at odds with this possibility. Nuclear pA^+ RNA foci are also observed in *S. cerevisiae* cells carrying a mutant allele of the MTR4 gene (Kadowaki et al., 1994). It is, therefore, an outstanding task, across different species, to discern whether macromolecular interactions exist, that are difficult to detect by conventional microscopy, and which coalesce into detectable foci only when the nuclear concentration of RNA increases upon exosome inactivation. Additionally, we do not have evidence to discriminate whether active RNA decay takes place in such putative RNP assemblies. Finally, *S. cerevisiae* strains debilitated in mRNA nuclear export also retain pA^+ RNA in nuclear focal structures (Hilleren et al., 2001; Jensen et al., 2001a; 2001b; Libri et al., 2002). Curiously, this retention is alleviated upon inactivation of the *S. cerevisiae* nuclear exosome, revealing that decay factors can also, in some situations, control nuclear RNA retention. Clearly, much still has to be learned about the relationships between nuclear RNA metabolism, localization, and export.

Although the mechanisms underlying the formation of RNA-containing foci are generally also not understood, a number of recent studies have demonstrated that non-membranous cellular compartments can form by so-called liquid-liquid phase transitions (Shin and Brangwynne, 2017). Their formation is often driven by sequence-specific physicochemical properties of RNA, facilitating multivalent protein-RNA and protein-protein interactions enriched with amino acid sequences of low complexity (Jain and Vale, 2017; Jiang et al., 2015; Kato et al., 2012; Uversky, 2017). The N-terminal part of ZFC3H1 harbors several low-complexity regions and is highly disordered (data not shown; using the low-complexity-predicting program PONDR-FIT; Xue et al., 2010), with particularly serine- and proline-rich stretches predicted to be important for phase separation (Shin and Brangwynne, 2017). Thus, it is likely that this property of ZFC3H1 drives formation of pA^+ RNA foci.

Consistent with their ZFC3H1 dependency, we found PAXT components, and the bona fide PAXT target SNHG19, to be enriched in pA^+ RNA foci. However, despite the absence of NEXT components ZCCHC8 and RBM7, we somewhat surprisingly also found the accumulation of NEXT substrates. Since PAXT complex loading onto target RNA presumably is pA-tail dependent (Meola et al., 2016), it is possible that NEXT substrates, stabilized in RNA exosome-depleted conditions, achieve pA-tails long enough to accommodate PAXT binding. Indeed, recent RNA-seq data (Meola et al., 2016) showed a partial stabilization of tested NEXT targets in ZFC3H1-depleted HeLa cells (see also Figures S3A and S3B), indicating that a fraction of transcripts can be degraded by both pathways. Moreover, NEXT targets tested in this study are highly expressed compared to bulk NEXT substrates and may, therefore, be especially prone to experiencing such pA tailing and subsequent loading of PAXT components. This suggestion is consistent with our previous observation that RBM7 binds to newly synthetized and unprocessed Pol II transcripts, including pre-mRNAs, and with the proposition that RBM7/NEXT recruits the exosome to trigger RNA degradation when unprotected RNA 3′ ends are available (Lubas et al., 2015). It is, therefore, possible that NEXT components, which are not present in pA^+ RNA foci, have already been discarded from these more processed/polyadenylated RNAs.

Our data imply that ZFC3H1-mediated nuclear retention of pA^+ RNA might occur in functional competition with its nuclear export by, e.g., AlyREF. Interestingly, AlyREF is loaded onto RNA in a 5′ cap-dependent manner (Cheng et al., 2006; Shi et al., 2017) and also binds mRNA 3′ ends PABPN1 dependently (Shi et al., 2017). It is conceivable that these interactions underlie aspects of transcript sorting between nuclear export and retention/degradation, with ZFC3H1 and AlyREF competing for transcript binding via a common interaction with PABPN1 as well as MTR4 competing with AlyREF for association with the CBC-bound ARS2 protein (Fan et al., 2017). Our analysis of AlyREF localization in exosome-depleted cells did not reveal any enrichment of this protein in pA^+ RNA foci. In contrast, ARS2, MTR4, and ZFC3H1 were all clearly enriched, indicating that accumulated transcripts are not immediately prone for export. It was proposed that formation of ARS2-MTR4 versus ARS2-AlyREF links is decisive in whether transcript gets degraded or exported, respectively (Fan et al., 2017). However, as previously reported (Andersen et al., 2013; Ogami et al., 2017), and as recapitulated here, MTR4 depletion in HeLa cells negatively affects the levels of ZFC3H1, ZCCHC8, and RRP40. Hence, it is difficult to address any direct impact of MTR4 on pA^+ RNA retention. Yet, RNA exosome targets escape the nucleus upon ZFC3H1 depletion (data in the present paper; Ogami et al., 2017), where MTR4 levels are unchanged, arguing against a direct role of MTR4 in transcript retention. It thus appears likely that PABPN1-dependent AlyREF versus ZFC3H1 loading on transcript 3′ends is a major RNA-sorting determinant, which also makes intuitive sense, since pA-tailing is a final step in RNA production, warranting an ultimate decision of whether to undergo export or decay. Moreover, it is possible that protein interactions with transcript 5′ and 3′ ends are cooperative so that any ZFC3H1-AlyREF 3′ end competition might affect ARS2-MTR4 5′ cap interactions and vice versa. Such cooperative factor loading was recently demonstrated by the depletion of CBP80, which impacts the 3′ end loading of AlyREF, and the depletion of the 3′ end processing factor CstF64, which disturbs 5′ cap association of AlyREF (Shi et al., 2017).

Our results are in agreement with a recent study reporting that the absence of ZFC3H1 results in the association of different types of pervasive transcripts with active ribosomes, causing global translation repression (Ogami et al., 2017). This positions ZFC3H1 as a key factor in preventing the nuclear export of RNAs that are seemingly unwanted in the cytoplasm. However, we also found that a fraction of fully spliced and full-length mRNAs was subject to exosomal decay. This suggests that bona fide translation substrates might also be targeted for nuclear turnover, echoing a recent demonstration that different cell types and tissues harbor 13%–30% of nuclear retained mRNAs (Halpern et al., 2015). Moreover, using annotation-agnostic classification of RNAs, Mukherjee et al. also found that certain transcription-factor-encoding transcripts were nuclear labile, and it was proposed that these RNAs may emanate from transcriptional bursts and that their nuclear retention might help buffer cytoplasmic transcript levels (Mukherjee et al., 2017). Consistently, this class of mRNA was upregulated upon RRP40 depletion (Figures 6D and S6G). Taken together, these data argue that nuclear turnover of full-length mRNA occurs. However, it remains an open question whether ZFC3H1-bound RNAs are committed to nuclear decay or whether the RNP assemblies, which we visualize as pA^+ RNA foci by inactivating the RNA exosome, might also serve a storage function capable of releasing, e.g., retained mRNA, in resemblance to cytoplasmic processing bodies (P-bodies) (Buchan, 2014). Previously reported nuclear retained mRNAs were localized in nuclear speckles (Halpern et al., 2015), the function of which is connected to active transcription (Galganski et al., 2017). Since pA^+ RNA foci do not co-localize with known nuclear bodies, including nuclear speckles, we suggest that these represent a post-transcriptional event, which is further supported by our observation that full-length spliced mRNA may appear in pA^+ RNA foci. Future research will be focused on determining how sorting between productive export and retention/decay is orchestrated and which additional factors contribute to this process.

EXPERIMENTAL PROCEDURES

Cell Culture and Transfections

HeLa cells were grown in DMEM containing 10% fetal bovine serum and 1% penicillin/streptomycin at 37°C and 5% CO_2. Cells were seeded at low confluence (approximately 44 cells per square millimeter) 1 day before siRNA transfection. The following day, DMEM was replaced with DMEM without penicillin/streptomycin, and transfections were performed using Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's instructions, with a final concentration of 20 nM of each siRNA (Table S1) and Lipofectamine 2000 (final dilution, 1:1,000) in RPMI 1640 medium. Two days after siRNA transfection, the medium was changed to new DMEM without penicillin/streptomycin, and the transfection procedure was repeated. For our complementation experiment, the ZFC3H1 coding construct pcDNA5-ZFC3H1-3xFLAG (Meola et al., 2016) or its empty vector pcDNA5-3xFLAG was transfected (using Lipofectamine 2000) into cells 1 day after the first siRNA transfection, and cells were harvested 1 day after the second siRNA transfection.

RNA-FISH Analysis

Cells were grown on microscope coverslips placed in tissue culture dishes. After siRNA treatment, medium was removed, and cells were washed twice with PBS prior to fixation with 4% formaldehyde in PBS for 20 min at room temperature. After fixation, cells were washed twice with PBS and permeabilized overnight at −20°C in 70% ethanol. Subsequently, cells were washed twice for 5 min with 2× SSC and blocked for 15 min in 2× SSC/30% formamide solution. The relevant FISH probe was added to hybridization mixture 1 (30% formamide, 2× SSC, 1 μg/μL *E. coli* tRNA), heated for 5 min at 95°C, and then cooled on ice. Thereafter, hybridization mixture 1 was combined with hybridization mixture 2 (10% dextran, 0.2 mg/mL BSA, 40 U RiboLock RNase inhibitors [Thermo Fisher Scientific]) and placed on a 37°C thermostat. 20 μL drops (per 12-mm coverslip) of combined hybridization mixtures 1 and 2 were pipetted onto a parafilm-coated box placed on wet filter paper, and coverslips were placed on the hybridization mixtures (cells facing the drops). The box was sealed with thin household film, and hybridization was performed overnight at 37°C. On the next day, coverslips were washed twice for 30 min each at 37°C in 2× SSC/30% formamide and twice for 5 min with PBS at room temperature. Cells were mounted with ProLong Gold Antifade Mountant with DAPI (Thermo Fisher Scientific).

ACTB, NFKB2, proDNAJB4, and RNVU1-14 3′ extension probes were designed using the R script Oligostan (Tsanov et al., 2016), and RNA-FISH was performed according to the smiFISH protocol described in Tsanov et al. (2016). All RNA-FISH probes used are listed in Table S2, except the oligo(dT)-LNA probe described in Thomsen et al., (2005).

Sequential RNA-FISH and IL Analysis

For co-staining of RNA and protein, RNA-FISH reactions were followed by protein IF staining. Briefly, a primary antibody dilution in 2% BSA was added to cells for 1 hr at room temperature, followed by three short washes with PBS. Thereafter, Alexa-Fluor-488-conjugated secondary antibody (Thermo Fisher Scientific) in 2% BSA was added for 1 hr at room temperature, followed by three short washes with PBS. Cells were mounted with ProLong Gold Antifade Mountant with DAPI. All primary antibodies and applied concentrations are listed in Table S3.

Microscopy and Image Analysis

Non-confocal images were obtained using a Zeiss Axiovert 200M epifluorescent microscope equipped with a coolSNAPHQ camera (Photometrics) and 60× (1.4) and 100× (1.3) objectives. Confocal images were obtained with a Zeiss LSM 710 confocal microscope and a 60× (1.4) objective. All images within the same experiment were taken with same excitation power and exposure time and processed similarly using the publicly available Fiji software (Schindelin et al., 2012). For all transcript-specific RNA-FISH images, background subtraction was performed using the Fiji Subtract Background plugin, with a rolling ball radius of 10 pixels. When image color balance adjustment was needed, all comparable images were processed in the same manner.

Western Blotting Analysis

Western blotting was performed by standard procedures (see Supplemental Experimental Procedures).

RNA Isolation, qRT-PCR Analysis, and Cellular Fractionation

RNA was purified using TRIzol reagent (Thermo Fisher Scientific) according to the manufacturer's instructions, and RNA was analyzed by qRT-PCR using standard procedures (see Supplemental Experimental Procedures).

RNA-Seq Library Preparation, Data Processing, and Computational Analysis

RNA-seq libraries were produced, and reads were quality controlled and mapped using standard procedures (see Supplemental Experimental Procedures).

Statistical Analysis

Values of biological replicates, p values, and statistical tests are reported in the figure legends. qRT-PCR data are shown as means ± SDs, except in Figure S6H where data are shown as means ± SEM. Statistical tests were done in the environment of the R Project for Statistical Computing (https://www.r-project.org).

DATA AND SOFTWARE AVAILABILITY

The accession number for the RNA-seq data reported in this paper is GEO: GSE108197.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and five tables and can be found with this article online at https://doi.org/10.1016/j.celrep.2018.04.061.

ACKNOWLEDGMENTS

This work was supported by the ERC (grant 339953), the Lundbeck Foundation, and the Novo Nordisk Foundation. We thank Marie Vestergård and Claudia Scheffler for technical support and Manfred Schmid for critical comments.

AUTHOR CONTRIBUTIONS

T.S. and T.H.J. conceived the project. T.S. designed, conducted, and analyzed the experiments. E.K. performed the bioinformatics analysis. M.L. and D.M. contributed to the RNA-FISH analysis. T.H.J. supervised the project. T.S. and T.H.J. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare that they have no competing interests.

Received: January 17, 2018
Revised: March 28, 2018
Accepted: April 15, 2018
Published: May 15, 2018

REFERENCES

Andersen, P.R., Domanski, M., Kristiansen, M.S., Storvall, H., Ntini, E., Verheggen, C., Schein, A., Bunkenborg, J., Poser, I., Hallais, M., et al. (2013). The human cap-binding complex is functionally connected to the nuclear RNA exosome. Nat. Struct. Mol. Biol. *20*, 1367–1376.

Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al. (2014). An atlas of active enhancers across human cell types and tissues. Nature *507*, 455–461.

Boulon, S., Verheggen, C., Jady, B.E., Girard, C., Pescia, C., Paul, C., Ospina, J.K., Kiss, T., Matera, A.G., Bordonné, R., and Bertrand, E. (2004). PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli. Mol. Cell *16*, 777–787.

Bresson, S.M., and Conrad, N.K. (2013). The human nuclear poly(A)-binding protein promotes RNA hyperadenylation and decay. PLoS Genet. *9*, e1003893.

Buchan, J.R. (2014). mRNP granules. Assembly, function, and connections with disease. RNA Biol. *11*, 1019–1030.

Chen, H.-M., Futcher, B., and Leatherwood, J. (2011). The fission yeast RNA binding protein Mmi1 regulates meiotic genes by controlling intron specific splicing and polyadenylation coupled RNA turnover. PLoS ONE *6*, e26804–e26815.

Cheng, H., Dufu, K., Lee, C.-S., Hsu, J.L., Dias, A., and Reed, R. (2006). Human mRNA export machinery recruited to the 5′ end of mRNA. Cell *127*, 1389–1400.

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012). Landscape of transcription in human cells. Nature *489*, 101–108.

Fan, J., Kuai, B., Wu, G., Wu, X., Chi, B., Wang, L., Wang, K., Shi, Z., Zhang, H., Chen, S., et al. (2017). Exosome cofactor hMTR4 competes with export adaptor ALYREF to ensure balanced nuclear RNA pools for degradation and export. EMBO J. *36*, 2870–2886.

Galganski, L., Urbanek, M.O., and Krzyzosiak, W.J. (2017). Nuclear speckles: molecular organization, biological function and role in disease. Nucleic Acids Res. *45*, 10350–10368.

Giacometti, S., Benbahouche, N.E.H., Domanski, M., Robert, M.C., Meola, N., Lubas, M., Bukenborg, J., Andersen, J.S., Schulze, W.M., Verheggen, C., et al. (2017). Mutually exclusive CBC-containing complexes contribute to RNA fate. Cell Rep. *18*, 2635–2650.

Hallais, M., Pontvianne, F., Andersen, P.R., Clerici, M., Lener, D., Benbahouche, Nel.H., Gostan, T., Vandermoere, F., Robert, M.C., Cusack, S., et al. (2013). CBC-ARS2 stimulates 3′-end maturation of multiple RNA families and favors cap-proximal processing. Nat. Struct. Mol. Biol. *20*, 1358–1366.

Halpern, K.B., Caspi, I., Lemze, D., Levy, M., Landen, S., Elinav, E., Ulitsky, I., and Itzkovitz, S. (2015). Nuclear retention of mRNA in mammalian tissues. Cell Rep. *13*, 1–11.

Harigaya, Y., Tanaka, H., Yamanaka, S., Tanaka, K., Watanabe, Y., Tsutsumi, C., Chikashige, Y., Hiraoka, Y., Yamashita, A., and Yamamoto, M. (2006). Selective elimination of messenger RNA prevents an incidence of untimely meiosis. Nature *442*, 45–50.

Hilleren, P., McCarthy, T., Rosbash, M., Parker, R., and Jensen, T.H. (2001). Quality control of mRNA 3′-end processing is linked to the nuclear exosome. Nature *413*, 538–542.

Hiriart, E., Vavasseur, A., Touat-Todeschini, L., Yamashita, A., Gilquin, B., Lambert, E., Perot, J., Shichino, Y., Nazaret, N., Boyault, C., et al. (2012). Mmi1 RNA surveillance machinery directs RNAi complex RITS to specific meiotic genes in fission yeast. EMBO J. *31*, 2296–2308.

Hrossova, D., Sikorsky, T., Potesil, D., Bartosovic, M., Pasulka, J., Zdrahal, Z., Stefl, R., and Vanacova, S. (2015). RBM7 subunit of the NEXT complex binds U-rich sequences and targets 3′-end extended forms of snRNAs. Nucleic Acids Res. *43*, 4236–4248.

Iasillo, C., Schmid, M., Yahia, Y., Maqbool, M.A., Descostes, N., Karadoulama, E., Bertrand, E., Andrau, J.-C., and Jensen, T.H. (2017). ARS2 is a general suppressor of pervasive transcription. Nucleic Acids Res. *45*, 10229–10241.

Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I.W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell *78*, 657–668.

Jain, A., and Vale, R.D. (2017). RNA phase transitions in repeat expansion disorders. Nature *546*, 243–247.

Jensen, T.H., Boulay, J., Rosbash, M., and Libri, D. (2001a). The DECD box putative ATPase Sub2p is an early mRNA export factor. Curr. Biol. *11*, 1711–1715.

Jensen, T.H., Patricio, K., McCarthy, T., and Rosbash, M. (2001b). A block to mRNA nuclear export in *S. cerevisiae* leads to hyperadenylation of transcripts that accumulate at the site of transcription. Mol. Cell *7*, 887–898.

Jiang, H., Wang, S., Huang, Y., He, X., Cui, H., Zhu, X., and Zheng, Y. (2015). Phase transition of spindle-associated protein regulate spindle apparatus assembly. Cell *163*, 108–122.

Jiang, L., 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-binding proteins and the Microprocessor to globally enhance pri-miRNA processing. Nat. Struct. Mol. Biol. *24*, 816–824.

Johnson, S.J., and Jackson, R.N. (2013). Ski2-like RNA helicase structures: common themes and complex assemblies. RNA Biol. *10*, 33–43.

Kadowaki, T., Chen, S., Hitomi, M., Jacobs, E., Kumagai, C., Liang, S., Schneiter, R., Singleton, D., Wisniewska, J., and Tartakoff, A.M. (1994). Isolation and characterization of *Saccharomyces cerevisiae* mRNA transport-defective (mtr) mutants. J. Cell Biol. *126*, 649–659.

Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J., et al. (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell *149*, 753–767.

Kilchert, C., Wittmann, S., and Vasiljeva, L. (2016). The regulation and functions of the nuclear RNA exosome complex. Nat. Rev. Mol. Cell Biol. *17*, 227–239.

Kung, J.T.Y., Colognori, D., and Lee, J.T. (2013). Long noncoding RNAs: past, present, and future. Genetics *193*, 651–669.

Libri, D., Dower, K., Boulay, J., Thomsen, R., Rosbash, M., and Jensen, T.H. (2002). Interactions between mRNA export commitment, 3′-end quality control, and nuclear degradation. Mol. Cell. Biol. *22*, 8254–8266.

Lubas, M., Christensen, M.S., Kristiansen, M.S., Domanski, M., Falkenby, L.G., Lykke-Andersen, S., Andersen, J.S., Dziembowski, A., and Jensen, T.H. (2011). Interaction profiling identifies the human nuclear exosome targeting complex. Mol. Cell *43*, 624–637.

Lubas, M., Andersen, P.R., Schein, A., Dziembowski, A., Kudla, G., and Jensen, T.H. (2015). The human nuclear exosome targeting complex is loaded onto newly synthesized RNA to direct early ribonucleolysis. Cell Rep. *10*, 178–192.

Marayati, B.F., Hoskins, V., Boger, R.W., Tucker, J.F., Fishman, E.S., Bray, A.S., and Zhang, K. (2016). The fission yeast MTREC and EJC orthologs ensure the maturation of meiotic transcripts during meiosis. RNA *22*, 1349–1359.

Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N., and Reed, R. (2005). Recruitment of the human TREX complex to mRNA during splicing. Genes Dev. *19*, 1512–1517.

Meola, N., and Jensen, T.H. (2017). Targeting the nuclear RNA exosome: poly(A) binding proteins enter the stage. RNA Biol. *14*, 820–826.

Meola, N., Domanski, M., Karadoulama, E., Chen, Y., Gentil, C., Pultz, D., Vitting-Seerup, K., Lykke-Andersen, S., Andersen, J.S., Sandelin, A., and Jensen, T.H. (2016). Identification of a nuclear exosome decay pathway for processed transcripts. Mol. Cell *64*, 520–533.

Morchikh, M., Cribier, A., Raffel, R., Amraoui, S., Cau, J., Severac, D., Dubois, E., Schwartz, O., Bennasser, Y., and Benkirane, M. (2017). HEXIM1 and NEAT1 long non-coding RNA form a multi-subunit complex that regulates DNA-mediated innate immune response. Mol. Cell *67*, 387–399.e5.

Moteki, S., and Price, D. (2002). Functional coupling of capping and transcription of mRNA. Mol. Cell *10*, 599–609.

Mukherjee, N., Calviello, L., Hirsekorn, A., de Pretis, S., Pelizzola, M., and Ohler, U. (2017). Integrative classification of human coding and noncoding genes through RNA metabolism profiles. Nat. Struct. Mol. Biol. *24*, 86–96.

Naganuma, T., and Hirose, T. (2013). Paraspeckle formation during the biogenesis of long non-coding RNAs. RNA Biol. *10*, 456–461.

Nizami, Z., Deryusheva, S., and Gall, J.G. (2010). The Cajal body and histone locus body. Cold Spring Harb. Perspect. Biol. *2*, a000653.

Ntini, E., Järvelin, A.I., Bornholdt, J., Chen, Y., Boyd, M., Jørgensen, M., Andersson, R., Hoof, I., Schein, A., Andersen, P.R., et al. (2013). Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. Nat. Struct. Mol. Biol. *20*, 923–928.

Ogami, K., Richard, P., Chen, Y., Hoque, M., Li, W., Moresco, J.J., Yates, J.R., 3rd, Tian, B., and Manley, J.L. (2017). An Mtr4/ZFC3H1 complex facilitates turnover of unstable nuclear RNAs to prevent their cytoplasmic transport and global translational repression. Genes Dev. *31*, 1257–1271.

Ohno, M., Segref, A., Bachi, A., Wilm, M., and Mattaj, I.W. (2000). PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. Cell *101*, 187–198.

Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M.S., Mapendano, C.K., Schierup, M.H., and Jensen, T.H. (2008). RNA exosome depletion reveals transcription upstream of active human promoters. Science *322*, 1851–1854.

Preker, P., Almvig, K., Christensen, M.S., Valen, E., Mapendano, C.K., Sandelin, A., and Jensen, T.H. (2011). PROMoter uPstream Transcripts share characteristics with mRNAs and are produced upstream of all three major types of mammalian promoters. Nucleic Acids Res. *39*, 7179–7193.

Ramanathan, A., Robb, G.B., and Chan, S.-H. (2016). mRNA capping: biological functions and applications. Nucleic Acids Res. *44*, 7511–7526.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Schneider, C., and Tollervey, D. (2014). Looking into the barrel of the RNA exosome. Nat. Struct. Mol. Biol. *21*, 17–18.

Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. Science *322*, 1849–1851.

Shi, M., Zhang, H., Wu, X., He, Z., Wang, L., Yin, S., Tian, B., Li, G., and Cheng, H. (2017). ALYREF mainly binds to the 5′ and the 3′ regions of the mRNA in vivo. Nucleic Acids Res. *45*, 9640–9653.

Shin, Y., and Brangwynne, C.P. (2017). Liquid phase condensation in cell physiology and disease. Science *357*, eaaf4382.

Sleeman, J.E., and Trinkle-Mulcahy, L. (2014). Nuclear bodies: new insights into assembly/dynamics and disease relevance. Curr. Opin. Cell Biol. *28*, 76–83.

St-André, O., Lemieux, C., Perreault, A., Lackner, D.H., Bähler, J., and Bachand, F. (2010). Negative regulation of meiotic gene expression by the nuclear poly(A)-binding protein in fission yeast. J. Biol. Chem. *285*, 27859–27868.

Sugiyama, T., and Sugioka-Sugiyama, R. (2011). Red1 promotes the elimination of meiosis-specific mRNAs in vegetatively growing fission yeast. EMBO J. *30*, 1027–1039.

Thomsen, R., Nielsen, P.S., and Jensen, T.H. (2005). Dramatically improved RNA in situ hybridization signals using LNA-modified probes. RNA *11*, 1745–1748.

Tsanov, N., Samacoits, A., Chouaib, R., Traboulsi, A.-M., Gostan, T., Weber, C., Zimmer, C., Zibara, K., Walter, T., Peter, M., et al. (2016). smiFISH and FISH-quant - a flexible single RNA detection approach with super-resolution capability. Nucleic Acids Res. *44*, e165.

Uversky, V.N. (2017). Intrinsically disordered proteins in overcrowded milieu: membrane-less organelles, phase separation, and intrinsic disorder. Curr. Opin. Struct. Biol. *44*, 18–30.

Visa, N., Izaurralde, E., Ferreira, J., Daneholt, B., and Mattaj, I.W. (1996). A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during nuclear export. J. Cell Biol. *133*, 5–14.

West, J.A., Davis, C.P., Sunwoo, H., Simon, M.D., Sadreyev, R.I., Wang, P.I., Tolstorukov, M.Y., and Kingston, R.E. (2014). The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. Mol. Cell *55*, 791–802.

Williams, T., Ngo, L.H., and Wickramasinghe, V.O. (2018). Nuclear export of RNA: Different sizes, shapes and functions. Semin. Cell Dev. Biol. *75*, 70–77.

Xue, B., Dunbrack, R.L., Williams, R.W., Dunker, A.K., and Uversky, V.N. (2010). PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. Biochim. Biophys. Acta *1804*, 996–1010.

Yamanaka, S., Yamashita, A., Harigaya, Y., Iwata, R., and Yamamoto, M. (2010). Importance of polyadenylation in the selective elimination of meiotic mRNAs in growing *S. pombe* cells. EMBO J. *29*, 2173–2181.

Yamashita, A., Shichino, Y., Tanaka, H., Hiriart, E., Touat-Todeschini, L., Vavasseur, A., Ding, D.Q., Hiraoka, Y., Verdel, A., and Yamamoto, M. (2012). Hexanucleotide motifs mediate recruitment of the RNA elimination machinery to silent meiotic genes. Open Biol. *2*, 120014.

Yamashita, A., Takayama, T., Iwata, R., and Yamamoto, M. (2013). A novel factor Iss10 regulates Mmi1-mediated selective elimination of meiotic transcripts. Nucleic Acids Res. *41*, 9680–9687.

Zhou, Y., Zhu, J., Schermann, G., Ohle, C., Bendrin, K., Sugioka-Sugiyama, R., Sugiyama, T., and Fischer, T. (2015). The fission yeast MTREC complex targets CUTs and unspliced pre-mRNAs to the nuclear exosome. Nat. Commun. *6*, 7050.

Zinder, J.C., and Lima, C.D. (2017). Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. Genes Dev. *31*, 88–100.

Zofall, M., Yamanaka, S., Reyes-Turcu, F.E., Zhang, K., Rubin, C., and Grewal, S.I.S. (2012). RNA elimination machinery targeting meiotic mRNAs promotes facultative heterochromatin formation. Science *335*, 96–100.

Cell Reports, Volume *23*

Supplemental Information

The RNA Exosome Adaptor ZFC3H1 Functionally Competes with Nuclear Export Activity to Retain Target Transcripts

Toomas Silla, Evdoxia Karadoulama, Dawid Mąkosa, Michal Lubas, and Torben Heick Jensen

Figure S1. Related to Figure 1

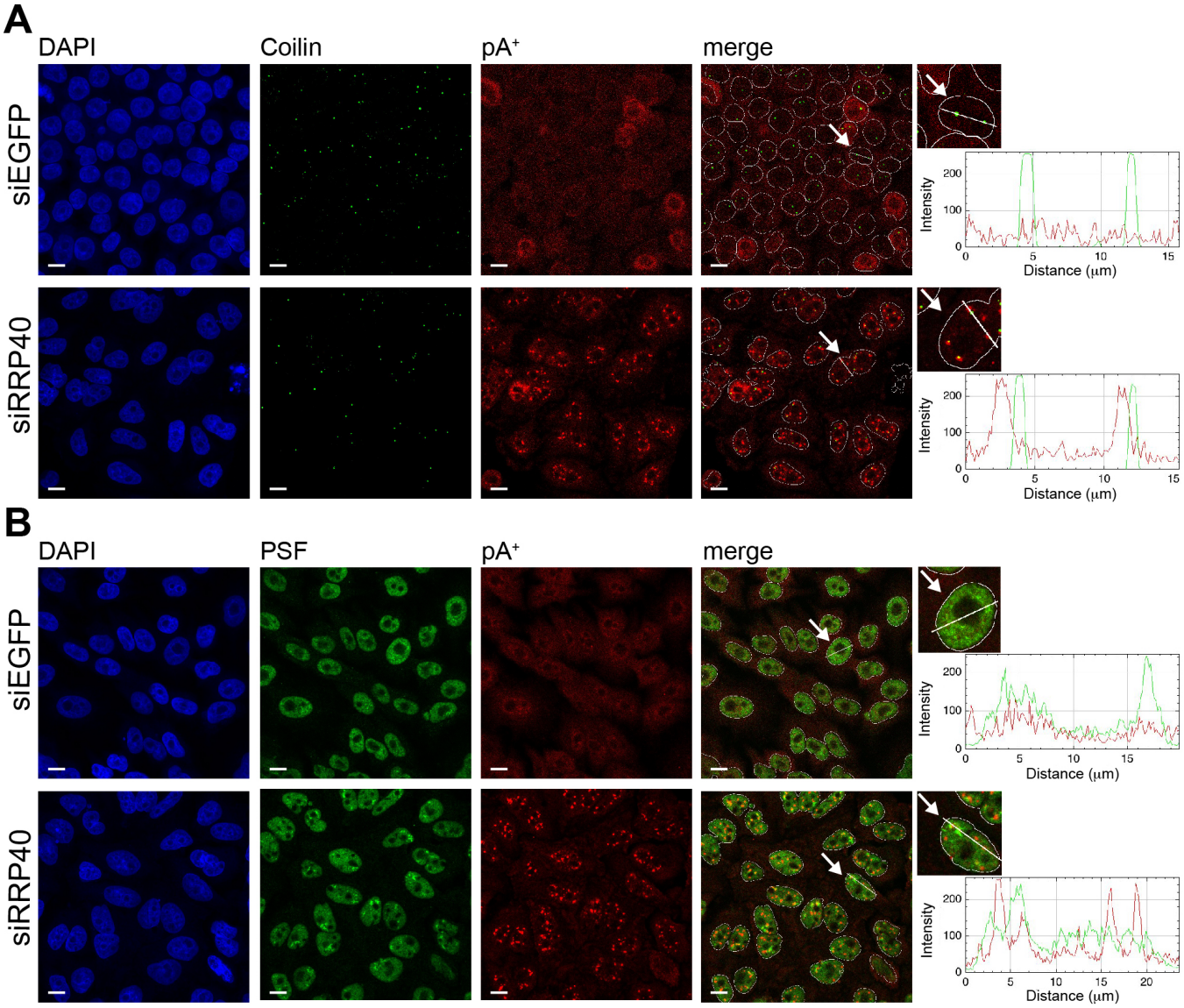


FIGURE S1. Related to Figure 1.
(A-B) Co-localization analysis of pA^+ RNA foci with the Cajal body marker Coilin (A) or the paraspeckle marker PSF (B). Images are displayed as in Figure 1C.

Figure S2. Related to Figure 2

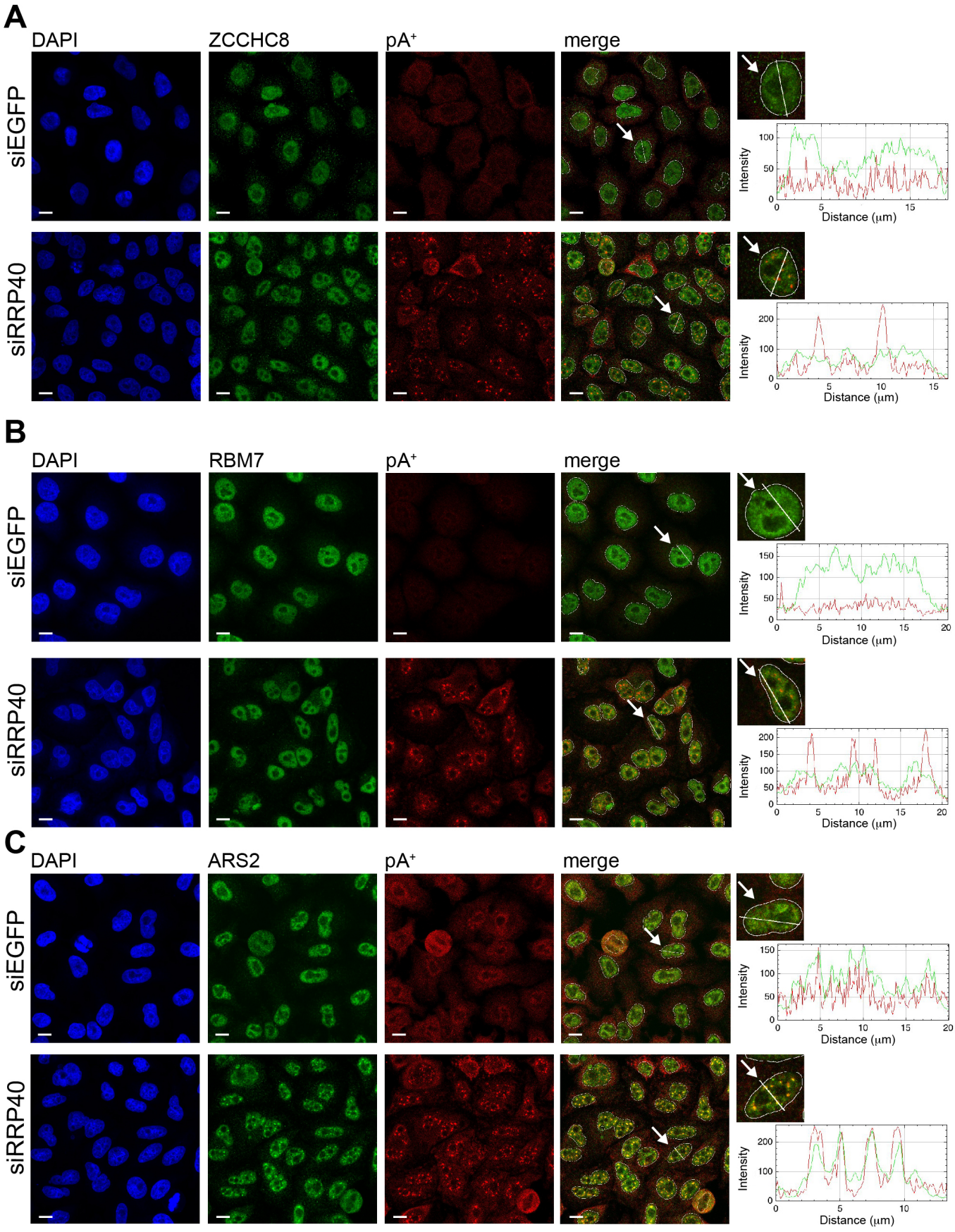


FIGURE S2. Related to Figure 2.
(A-C) Co-localization analysis of NEXT complex components ZCCHC8 (A) and RBM7 (B) as well as the CBCA component ARS2 (C) with pA$^+$ RNA in control (siEGFP) or RRP40 depleted (siRRP40) HeLa cells. Image display as in Figure 2.

Figure S3. Related to Figure 3

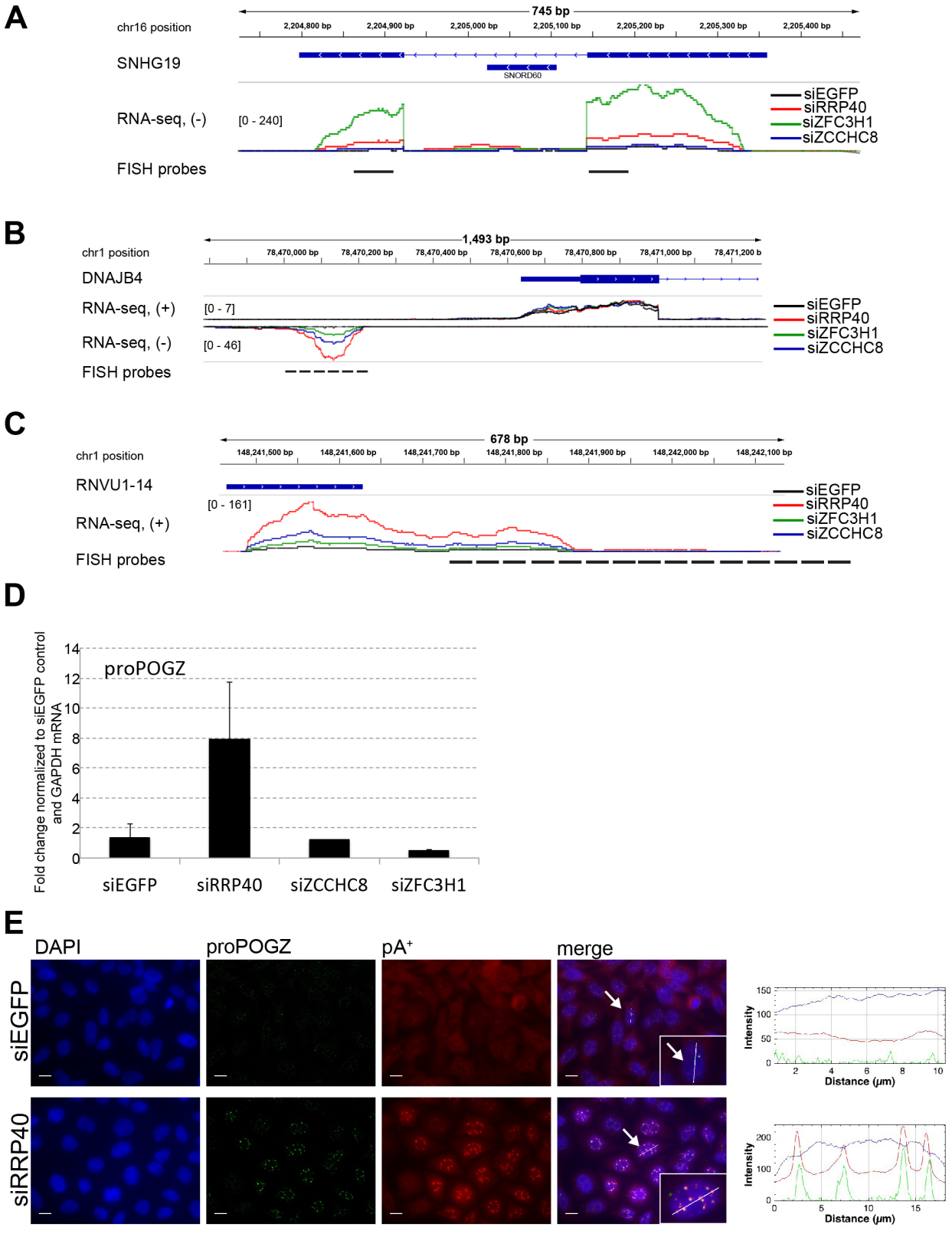


FIGURE S3. Related to Figure 3.
(A-C) Genome browser view of the SNHG19 (A), DNAJB4 (A) and RNVU1-14 (C) loci, showing tracks of normalized total RNA-seq reads from control (siEGFP, black line), RRP40 (red line), ZFC3H1 (green line) and ZCCHC8 (blue line) depleted HeLa cells (Meola et al., 2016). All tracks represent an average of three biological replicates. Solid black lines below the sequencing tracks indicate target sites of the used RNA FISH probes. Blue horizontal bars represent the Refseq annotation of the SNHG19 snoRNA host gene and the SNORD60 snoRNA (A), the 5` end of the DNAJB4 (B) and the RNVU1-14 (C) genes.
(D) RT-qPCR analysis of proDNAJB4, RNVU1-14 3`extension and exogenous proPOGZ transcript levels using total RNA harvested from control (EGFP), RRP40, ZCCHC8 and ZFC3H1, siRNA treated Hela cells as indicated. Data are displayed as mean values with error bars denoting SDs (n = 3 biological replicates).
(E) RNA FISH co-localization analyses of pA^+ RNA with exogenous proPOGZ transcript in control (siEGFP) or RRP40 (siRRP40) depleted HeLa cells. Image display as in Figure 3 A-C.

Figure S4. Related to Figure 4

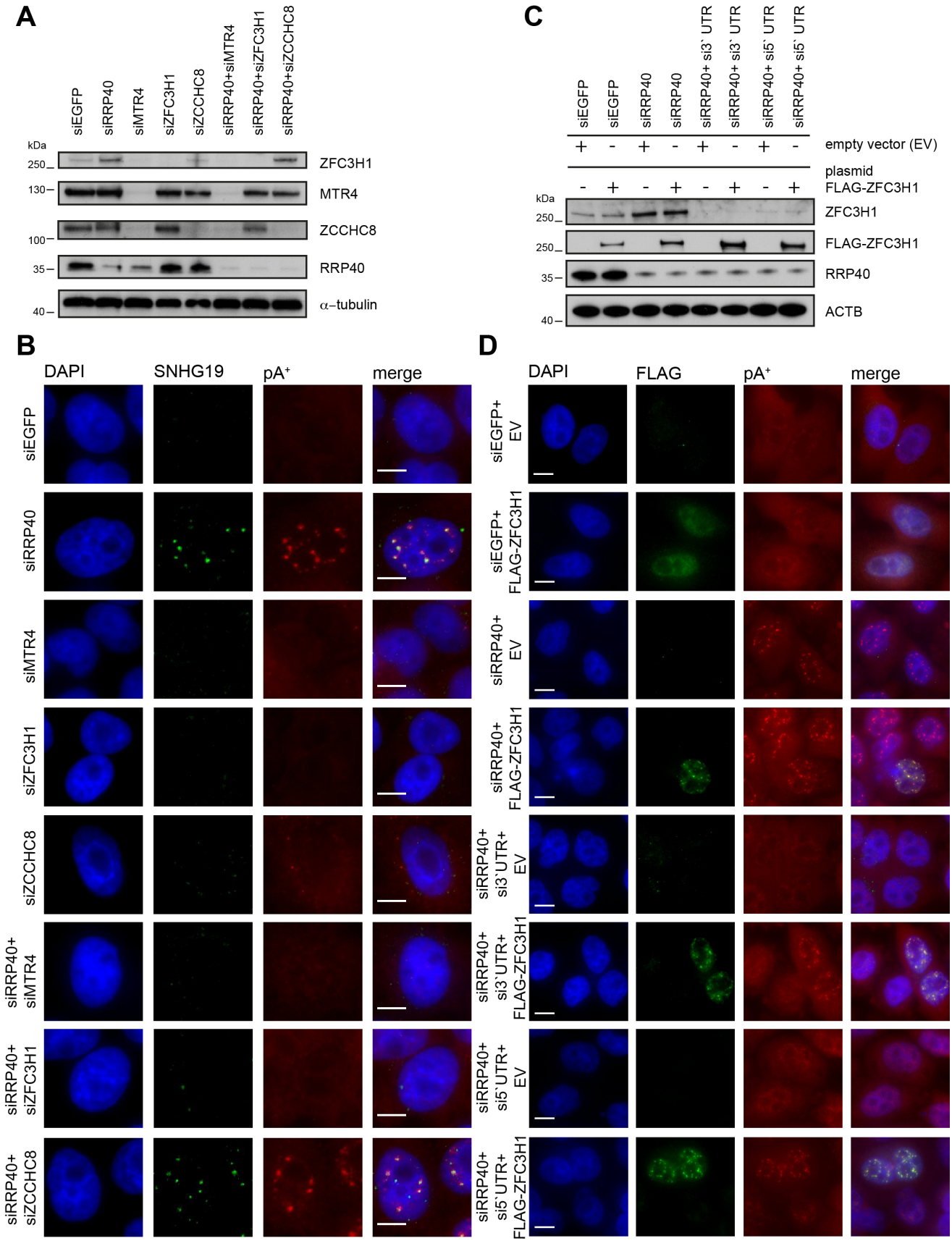


FIGURE S4. Related to Figure 4.
(A) Western blotting analysis using extracts from HeLa cells depleted for the indicated proteins. α-tubulin was used as loading control. Note that MTR4 depletion also affects ZFC3H1, ZCCHC8 and RRP40 protein levels.
(B) Dual SNHG19- and pA^+ RNA-FISH analysis from Figure 4A but with respective signals shown separately. pA^+ RNA was detected with an oligo(dT) 50-mer probe.
(C) Western blotting analysis using extracts from HeLa cells depleted for the indicated proteins and expressing FLAG-tagged exogenous ZFC3H1 or not as indicated. HeLa cells were co-transfected with the indicated siRNAs and the empty vector (EV) or FLAG-ZFC3H1 constructs. Beta-actin (ACTB) was used as loading control. si3'UTR and si5'UTR target the ZFC3H1 RNA 3' and 5' UTRs, respectively. Note that the FLAG-ZFC3H1 expression is not discernable using the ZFC3H1 antibody due to a low plasmid transfection efficiency.
(D) Parallel pA^+ RNA FISH analysis of HeLa cells from (C). Images of DAPI (blue), anti-FLAG antibody (green) and pA^+ RNA (red) are shown together with their merged images. pA^+ RNA was detected with an oligo(dT)-LNA probe. Note that and the pA^+ RNA signal is less exposed in panel (B) compared to (D). Scale bar: 10 μm.

Figure S5. Related to Figure 5

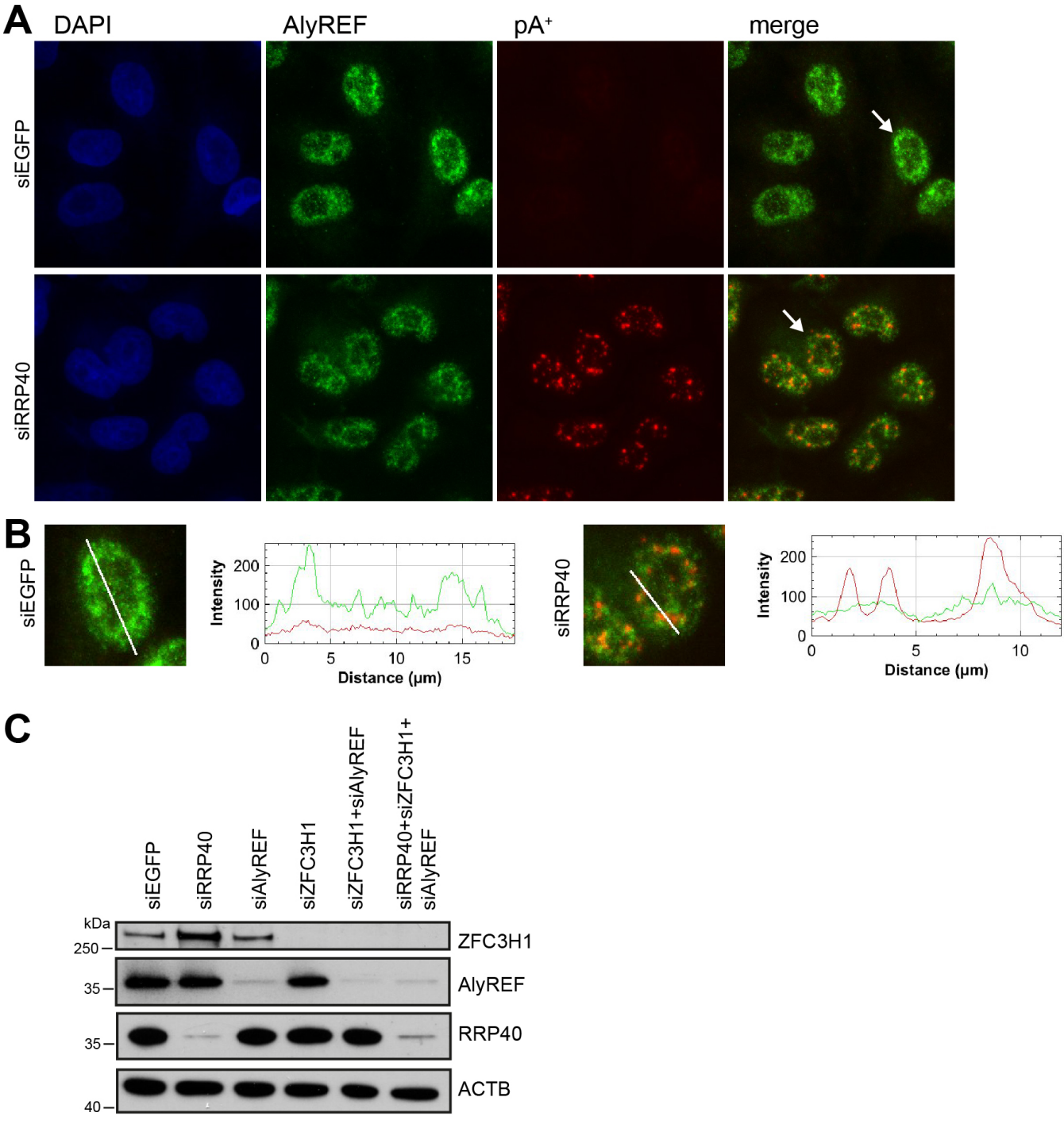


FIGURE S5. Related to Figure 5.
(A) Co-localization analysis of AlyREF and pA$^+$ RNA in control (siEGFP) or RRP40 (siRRP40) depleted HeLa cells. Image display as in Figure 1C.
(B) Zoom-in of cells pointed to by arrows in (A). Plot profiles as in Figure 1C.
(C) Western blotting analysis of levels of the indicated factors using extracts from HeLa cells depleted for the indicated proteins. ACTB (beta-actin) was used as loading control.

Figure S6. Related to Figure 6

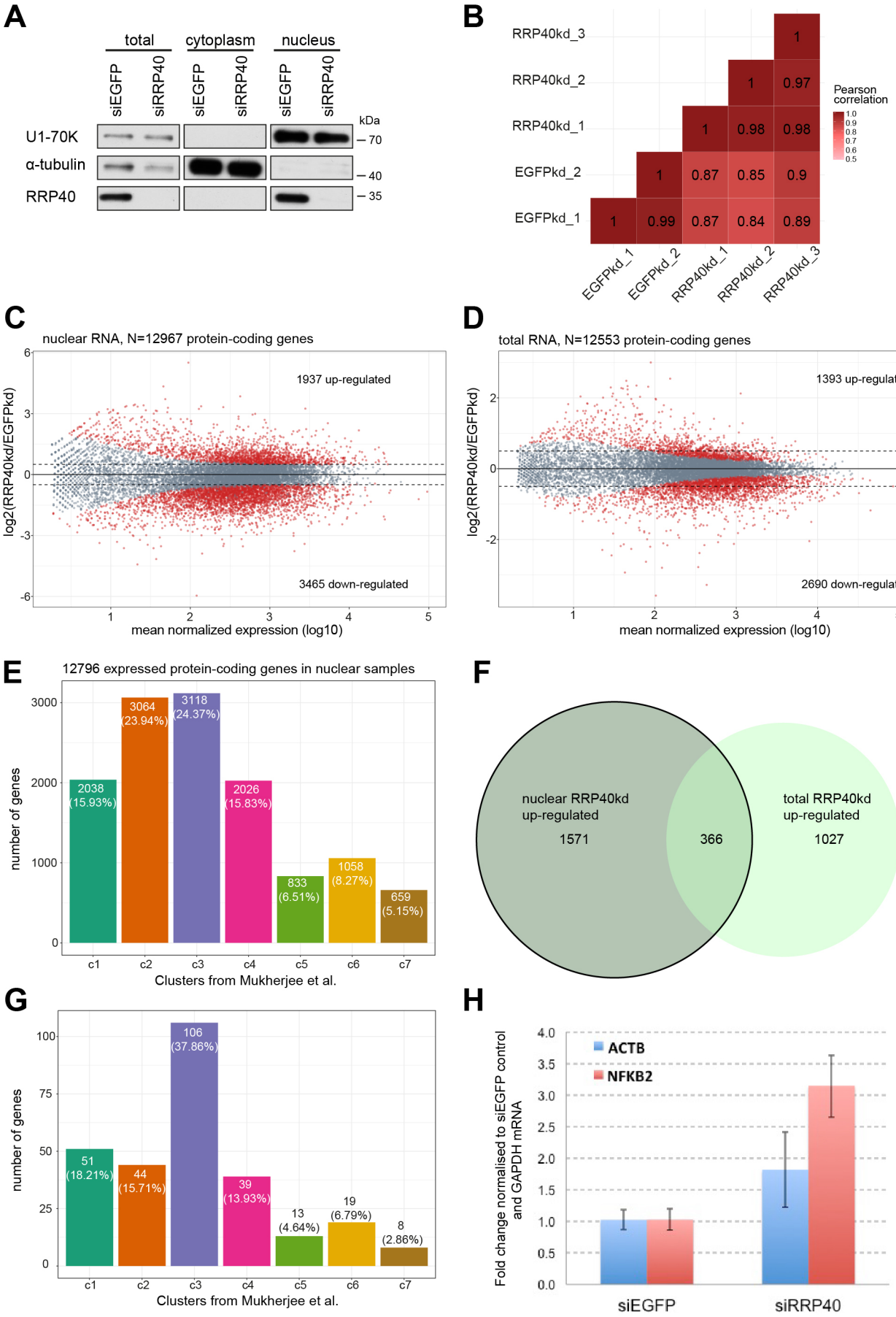


FIGURE S6. Related to Figure 6.

(A) Western blotting analysis of total, cytoplasmic and nuclear fractions extracted from control (siEGFP) and RRP40 (siRRP40) depleted HeLa cells. U1-70K and α-tubulin are used as nuclear and cytoplasmic markers, respectively. (B) Pearson correlation matrix of expression values of all expressed GENCODE annotated genes between the two control (EGFPkd) and three RRP40kd samples. (C and D) MAplot of all expressed protein-coding genes in nuclear (C) and total (D) RNAseq libraries. Results represent differential expression of RRP40kd compared to control samples. Statistically significant hits (padj<0.05) are color-coded red.

(E) Gene classification of all expressed genes in nuclear HeLa samples according to Mukherjee et.al (Mukherjee et al., 2016)

(F) Venn diagram displaying the overlap between significantly up-regulated (log2(RRP40kd/EGFPkd)>0, padj<0.05) protein coding genes in the nuclear and total RNA libraries.

(G) Bar plots showing the distribution of the overlapping genes from (E) according to a previously published coding and non-coding RNA classification (Mukherjee et al., 2016). Note the enrichment of genes belonging to cluster c3 (Pearson's Chi-squared test, p-value = 3.207e-07; against background of all expressed genes in nuclear samples shown in (E)).

(H) RT-qPCR analysis of ACTB and NFKB2 transcripts from nuclear RNA fractions expressed as ratios between RRP40-depleted and control HeLa cells. Data were displayed as mean values with error bars denoting s.e.m (n = 3 biological replicates).

Figure S7. Related to Figure 6

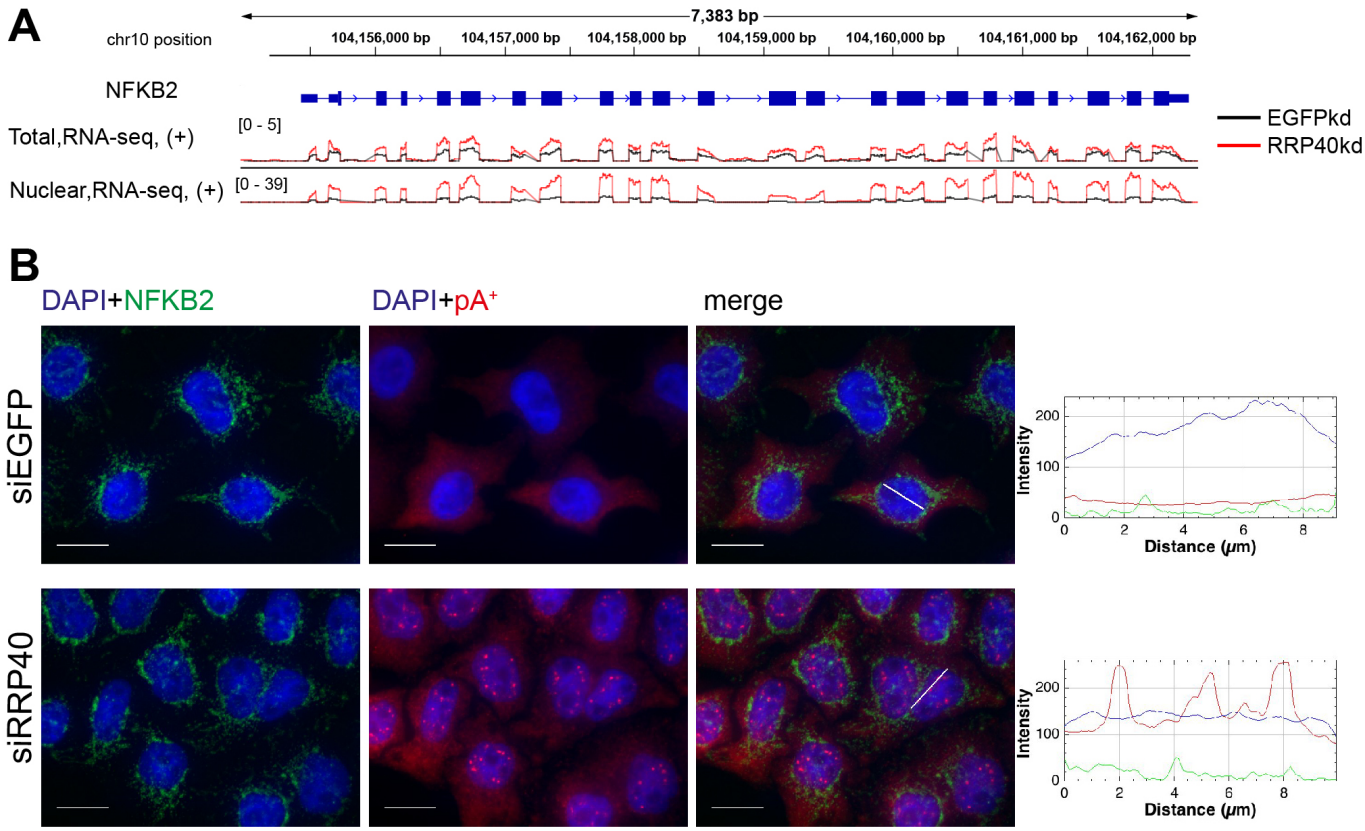


FIGURE S7. Related to Figure 6.
(A) Genome browser view of the NFKB2 locus, showing tracks of normalized total and nuclear RNA-seq data from control (siEGFP, black line) and RRP40 (red line) depleted HeLa cells. Tracks represent average of three biological replicates.
(B) Dual NFKB2- and pA^+ RNA-FISH analysis in the indicated factor-depleted HeLa cells. Image display as in Figure 1C.

Table S1. siRNA oligonucleotides. Related to Experimental Procedures

Target	Sense (5'-3')	Antisense (5'-3')
EGFP	GACGUAAACGGCCACAAGUdTdT	ACUUGUGGCCGUUUACGUCdTdT
ZCCHC8	GGAAUGUACCUCAGGAUAAdTdT	UUAUCCUGAGGUACAUUCCdTdT
ZFC3H1	GAUUAGAGUCCAUGAUUAAdTdT	UUAAUCAUGGACUCUAAUCdTdT
RRP40	CACGCACAGUACUAGGUCAdTdT	UGACCUAGUACUGUGCGUGdTdT
AlyREF	CGUGCACUUUGAGCGGAAGdTdT	CUUCCGCUCAAAGUGCACGdTdT
MTR4	CACGCACAGUACUAGGUCAdTdT	UGACCUAGUACUGUGCGUGdTdT
PABPN1	GUAGAGAAGCAGAUGAAUAdTdT	UAUUCAUCUGCUUCUCUACdTdT
ZFC3H1 5`UTR	GGGCUAAGGUUGUGUGGAAdTdT	UUCCACACAACCUUAGCCCdTdT
ZFC3H1 3`UTR	GGUCAAAUAUUAUGUGCAAdTdT	UUGCACAUAAUAUUUGACCdTdT

Table S2. FISH probes. Related to Experimental Procedures (Excel sheet)

Table S3. Antibodies used in sequential RNA-FISH and Immunofluorescence. Related to Experimental Procedures

Target	**Source**	**Cat # reference**	**Working dilution**
MTR4	Abcam	ab70551	1:1000
ZFC3H1	Novus Biologicals	NB100-68267	1:500
PABPN1	Abcam	ab75855	1:1000
ARS2	Abcam	ab88392	1:1000
ZCCHC8	Abcam	ab68739	1:500
RBM7	Sigma-Aldrich	HPA013993	1:1000
AlyREF	Sigma-Aldrich	HPA019799	1:1000
PSF	Sigma-Aldrich	P2860	1:1000
Coilin	Abcam	ab87913	1:1000
SC35	Abcam	ab11826	1:1000
FLAG	Sigma-Aldrich	F3165-2MG	1:2000

Table S4. Antibodies used in western experiments. Related to Experimental Procedures

Target	**Source**	**Cat # reference**	**Working dilution**
MTR4	Abcam	Ab70551	1:2500
ZFC3H1	Novus Biologicals	NB100-68267	1:1000
ZCCHC8	Abcam	Ab68739	1:2500
β-ACTIN	Sigma-Aldrich	A2228	1:100000
RRP40	Proteintech	15062–1-AP	1:4000
α-TUBULIN	Rockland	600-401-880	1:2500
U1-70K	custom made		1:10000
AlyREF	Abcam	Ab202894	1:10000
FLAG	Sigma-Aldrich	F3165-2MG	1:10000

Table S5. qPCR oligonucleotides. Related to Experimental Procedures

Target	Forward (5'-3')	Reverse (5'-3')
GAPDH	GTCAGCCGCATCTTCTTTTG	GCGCCCAATACGACCAAATC
SNHG19	CGTCCAGGCCTGGCCTAC	GCTCGCGACGAAACCTGC
RNVU1-14 3` ext	CTGTCCGGTCAGTTGTTTCC	GGAGAACAACCCAACCAACA
proPOGZ	GCTGGGCCTGCAAATAAATA	TTGGCAAGTAAATTGGGGAATA
proDNAJB4	TTTCTGGCGTTTCTGATTGA	ACCAAAACGCAGGTTGTTTA
ACTB mRNA	GAGACCGCGTCCGCC	ATCATCCATGGTGAGCTGGC
NFKB2 mRNA	TCTACTGGAGGCCCTGTCTG	TTCCTTCACCTCTGTGCTGG

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Western blotting analysis
Cell pellets were re-suspended in RSB100 (100 mM Tris, pH 7.5, 100 mM NaCl, 2.5 mM MgCl2, 0.5% Triton X-100) and incubated on ice for 10 min. Cell debris was removed by centrifugation at 4000g, 4°C for 15 min. Samples were separated by 10% denaturing PAGE and transferred to PVDF membranes (Millipore), which were blocked in 5% skimmed-milk powder (SMP) in PBS for one hour at room temperature. Primary antibody solutions in 5% SMP in PBS were added to the membranes and incubated one hour at room temperature. Membranes were then washed three times for 5 min in PBS/0.05% Tween 20 and incubated in horseradish peroxidase (HRP)-conjugated goat-anti-rabbit or -mouse secondary antibody (Dako) in 5% SMP in PBS, washed again three times for 5 min in PBS/0.05% Tween 20 and exposed using Supersignal West Femto Substrate (Thermo Fischer Scientific). All used antibodies and applied concentrations are listed in Table S4.

Cellular fractionation
Cells were collected in PBS/3mM EDTA, pelleted by centrifugation and lysed on ice in 1 ml (per 150 mm dish) fractionation buffer (10 mM Tris/HCl [pH 7.4], 10 mM NaCl, 2.5 mM MgCl2, 1mM DTT, 200 µg/ml Digitonin, protease inhibitors [Roche], 40 U RiboLock RNase inhibitors [Thermo Fisher Scientific]) by gentle pipetting. Lysates were left on a tube rotator for 10 min at 4°C. Subsequently, aliquots were taken for total RNA and protein analysis. Nuclei were pelleted by centrifugation at ~1000g for 5 min at 4°C. After centrifugation the supernatant (cytoplasmic fraction) was transferred to a new tube and centrifuged at maximum speed in a tabletop centrifuge at 4°C. The nuclear pellet was washed in 1 ml of homogenization buffer (10 mM Tris/HCl [pH 7.4], 15 mM NaCl, 60 mM KCl, 150mM Sucrose, 0.3% NP40) and centrifuged at ~1000x g for 5 min at 4°C. The nuclear pellet was re-suspended in 1 ml of S1 solution (0.25 M sucrose, 10 mM MgCl2) and added on top of an equal volume of S3 solution (0.88 M sucrose, 0.5 mM MgCl2), and centrifuged at 2800g for 10 min at 4°C. Supernatant was discarded and the nuclear pellet was re-suspend in 1 ml of wash buffer (10 mM Tris/HCl [pH 7.4], 15 mM NaCl, 60 mM KCl) and centrifuged at 1000x g for 5 min. After centrifugation the supernatant was discarded and the nuclear pellet was washed with 1 ml PBS and centrifuged again at 1000x g for 5 min. The resulting supernatant was discarded and the nuclear pellet was re-suspended in PBS and used for RNA extraction and protein analysis.

RNA isolation and RT-qPCR analysis
RNA was purified using TRIzol reagent (Thermo Fisher Scientific) according to the manufacturer's instructions. Purified RNA was DNase I treated (TURBO DNA-free Kit, Thermo Fisher Scientific). For RT-qPCR analysis cDNA was synthesized using random hexamer primers according to standard procedures. qPCR was performed using primers from Table S5 on a AriaMx Real-time qPCR System (Agilent). Data were processed using the ΔΔCt method, with normalization to both GAPDH mRNA levels and EGFP-siRNA control samples.

RNAseq library preparation and data processing
50 µg of nuclear RNA was used for pA^+ RNA purification using Oligo (dT)25 Dynabeads (Thermo Fisher Scientific) according to the manufacturer's instructions. ~750 ng of nuclear pA^+ RNA was treated with Ribo-ZeroTM Magnetic Gold Kit (Human/Mouse/Rat; Illumina) according to the manual. cDNAs were produced using random hexamer-primers. Strand specific libraries and their sequencing on Illumina HiSeq 2000 were performed at BGI Tech Solutions (Hong Kong, China). Sequence reads were quality controlled using FastQC v0.11.2 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Trimmomatic v0.32 was used to remove remainders of illumine adaptors, low quality bases, the first 12 bases of the reads and reads below 25 nucleotides. The settings used were "ILLUMINACLIP:<TrueSeq3_PE_2>:2:30:10 HEADCROP:12 LEADING:22 SLIDINGWINDOW:4:22 MINLEN:25" (Bolger et al., 2014). Trimmed reads were mapped against the human genome 19 (hg19), using HISAT v0.1.6.beta (Kim et al., 2015) by providing a list of GENCODE-annotated splice sites, setting the maximum fragment length to 1000 and using parameter --rf for the upstream/downstream mate orientation, and otherwise

using default settings. Samtools v0.1.17 (Li et al., 2009) were utilized to select uniquely mapped and properly paired reads, which were then used for downstream analysis.

Gene annotation
For annotation all genes from the human GENCODE annotation v19 (Harrow et al., 2012) were used together with PROMPT and eRNA sets from (Meola et al., 2016). Exons and introns were defined from the GENCODE annotation, that is, all nucleotides within a gene that could be annotated as exons in any of the isoforms of the gene, were annotated as exonic and all the remaining nucleotides within the gene region were annotated as intronic. R and python scripts were used and are available upon request.

RNAseq quantification
The Rsubread package (Liao et al., 2013) from the Bioconductor R package (http://www.bioconductor.org) was used to quantify expression in control (EGFPkd) and exosome depleted (RRP40kd) nuclear pA^+ and total libraries (GEO database ID: GSE84172, (Meola et al., 2016)) by counting the RNAseq fragments that overlapped exons on the relevant strand and summarized on the gene level. Fragment counts were converted to tags per million (where tags were defined as uniquely mapped and properly paired fragments). The quantification was done for all genes in the human GENCODE annotation v19, for PROMPTs and eRNAs used previously (Meola et al., 2016) and for all collapsed exons and introns defined using the GENCODE annotation v19. Then, for Figures 6A and S6C-D, the DESeq2 package (Love et al., 2014) was used to perform differential expression analysis between the control and the knock down samples for both total and nuclear $polyA^+$ RNAseq libraries. In Figure 6B, exon and intron exosome sensitivity was calculated based on the strand specific expression using following formula:

$$(1)\ \text{sensitivity}_{\text{KD}} = \frac{\text{expression}_{\text{KD library}} + \varepsilon}{\text{expression}_{\text{CTRL library}} + \varepsilon},$$

where the expression values are in TPMs and ε is a pseudocount defined as the minimum expression value of all RNAs in all libraries min(expression_ALL>0)

Statistics and RNAseq data visualisation
Visualisations were performed using the ggplot2 R package (Wickham et al, 2009). Statistical tests were done in the environment of the R Project for Statistical Computing (https://www.r-project.org).

SUPPLEMENTAL REFERENCES

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data . Bioinformatics 30, 2114–2120.

Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., et al. (2012). GENCODE: The reference human genome annotation for The ENCODE Project. Genome Res 22, 1760–1774.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat Meth 12, 357–360.

Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Research 41, e108–e108.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550–21.

Meola, N., Domanski, M., Karadoulama, E., Chen, Y., Gentil, C., Pultz, D., Vitting-Seerup, K., Lykke-Andersen, S., Andersen, J.S., Sandelin, A., et al. (2016). Identification of a Nuclear Exosome Decay Pathway for Processed Transcripts. Molecular Cell 64, 520–533.

Mukherjee, N., Calviello, L., Hirsekorn, A., de Pretis, S., Pelizzola, M., and Ohler, U. (2016). Integrative classification of human coding and noncoding genes through RNA metabolism profiles. Nat. Struct. Mol. Biol. 24, 86–96.

Wickham H (2009). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York