THE ROLE OF THE HUMAN NUCLEAR POLY(A) BINDING PROTEIN IN RNA DECAY

APPROVED BY SUPERVISORY COMMITTEE

Nicholas Conrad, PhD

Carla Green, PhD

Michael Buszczak, PhD

Qinghua Liu, PhD

THE ROLE OF THE HUMAN NUCLEAR POLY(A) BINDING PROTEIN IN RNA DECAY

by

STEFAN BRESSON

THESIS

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSPOPHY

The University of Texas Southwestern Medical Center

Dallas, Texas

May, 2015

THE ROLE OF THE HUMAN NUCLEAR POLY(A) BINDING PROTEIN IN RNA DECAY

Stefan Bresson

The University of Texas Southwestern Medical Center at Dallas, 2014

Nicholas Conrad, Ph.D.

Control of nuclear RNA stability is an important determinant of gene expression, but the factors involved in nuclear RNA decay remain largely unknown in higher eukaryotes. Here, I describe our work showing that polyadenosine (poly(A)) tails can stimulate transcript decay in the nucleus of human cells, a function mediated by the ubiquitous nuclear poly(A) binding protein PABPN1. We show that PABPN1 is required for the degradation of a viral nuclear noncoding RNA as well as an inefficiently exported human mRNA. Importantly, the targeting of RNAs to this decay pathway requires the PABPN1 and poly(A) polymerase (PAP)-dependent extension of the poly(A) tail. Nuclear transcripts with longer poly(A) tails are then selectively degraded by components of the nuclear exosome. I also describe our work showing that PABPN1 and PAP are required for the degradation of a variety of nuclear-retained long noncoding RNAs. Taken together, this work uncovers an important pathway in the turnover of RNAs in the nucleus.

Table of Contents

Prior Publications

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

List of Figures

List of Figures (cont.)

List of Abbreviations

4SU: 4-thiouridine

4-thioUTP: 4-thiouridine triphosphate

CHX: cycloheximide

CPSF: cleavage and polyadenylation specificity factor

DEG: differentially expressed gene

ENE: expression and nuclear retention element

KSHV: Kaposi's sarcoma associated herpesvirus

NMD: nonsense mediated decay

NTP: nucleotide triphosphate

PABPN1: poly(A) binding protein nuclear 1

PAN: polyadenylated nuclear RNA

PAP: poly(A) polymerase

PPD: PABPN1-PAP decay

QC: quality control

RPKM: reads per kilobase of gene per million reads

scaRNA: small cajal body RNA

snoRNA: small nucleolar RNA

SNHG: snoRNA host gene (noncoding)

TRAMP: Trf4/5-Air1/2-Mtr4p polyadenylation complex

Chapter 1: Literature Review

Prior to their export to the cytoplasm and use in translation, eukaryotic messenger RNAs (mRNAs) must undergo a series of maturation steps. mRNAs are transcribed as precursor RNAs consisting of both exons and introns. Introns are removed and degraded, while the intervening exons are spliced together to generate the mature mRNA. Additional modifications at the 5′ and 3′ ends of the RNA protect the message from decay. A 7-methylguanosine cap is added at the 5′ end, and a polyadenosine (poly(A)) tail is added at the 3′ end. Mistakes in these processes result in defective mRNAs which may code for aberrant proteins with dominant negative effects. Consequently, cells have evolved a variety of RNA surveillance pathways to identify and eliminate misprocessed RNAs. While these pathways have been well characterized in yeast, they remain largely unknown in higher eukaryotes.

Mammalian splicing

Eukaryotic pre-mRNAs consist of both introns and exons. In a process known as splicing, introns are removed and degraded, and exons are joined to form the mature, coding mRNA. Though splicing is universally conserved in eukaryotes, its frequency varies dramatically among species. In *S. cerevisae*, most genes are intronless (~95%), but the intron containing genes are highly expressed and approximately one-third of newly made mRNAs contain introns (Ares et al., 1999). In mammalian cells, intronless genes are the exception rather than the rule, as only ~5% of human mRNAs are intronless. Moreover, intronless genes tend to be tissue specific and expressed at low levels (Shabalina et al., 2010).

Splicing is thought to serve several useful functions. First, splicing allows the cell to mix and match different exons. Exons can be extended or omitted entirely, and introns can be retained. Thus, one gene can be used to generate multiple unique protein isoforms in a tissue or context-specific manner. In human cells, an estimated 95% of multi-exon genes undergo alternative splicing (Pan et al., 2008), but it is still unclear to what extent these represent useful products (i.e. beneficial to the cell). Second, introns can be used to encode functional products. For example, most known mammalian small nucleolar RNAs (snoRNAs) are encoded within introns, and processed following splicing. In most cases they are found within protein coding genes, while a small number are found within the introns of dedicated noncoding snoRNA host genes (Rearick et al., 2011). Third, splicing plays an important role in efficient gene expression. One of the first transcripts in which this was discovered was the human β-globin RNA. Mature β-globin levels were shown to be highly dependent on splicing: the intron-containing version is expressed ~400 times higher than the corresponding cDNA-derived transcript (Buchman and Berg, 1988). Importantly, insertion of a heterologous intron was able to rescue the poor expression of the β-globin cDNA, demonstrating that the expression differences were due to splicing, rather than the specific sequence of the intron itself. These results may be explained, at least in part, by the relatively inefficient export of intronless β-globin. Splicing increases the cytoplasmic/nuclear ratio of β-globin more than 10 fold (Valencia et al., 2008). Consequently, intronless β-globin is subject to a different set of RNA decay factors than its cytoplasmic counterpart. Since mRNA levels are determined by both synthesis and decay rates, differential stability could explain the different steady state levels. Consistent with this hypothesis, intronless β-globin is highly unstable compared to spliced β-globin. (Conrad et al., 2006). Importantly, the

factors involved in intronless β-globin decay, and mammalian nuclear RNA decay generally, remain unknown.

It is now clear that β-globin is an exceptional case, but most genes are still somewhat dependent on splicing for optimal expression. Depending on the particular gene, splicing has since been shown to affect essentially every aspect of gene expression, including transcription, 3′ end formation, export, stability, and translation (Moore and Proudfoot, 2009). Moreover, these effects are seen throughout eukaryotes (Callis et al., 1987; Goebels et al., 2013; Rose, 2004), suggesting that splicing has a conserved role in stimulating gene expression. As mentioned above, ~5% of human genes are intronless. These genes require the presence of specific cis-acting elements to promote export and/or nuclear RNA stability (Lei et al., 2011). Viral genes are also frequently intronless, and viruses have evolved multiple strategies to overcome splicing-dependent export and stability effects (Chen et al., 2002; Conrad and Steitz, 2005; Fischer et al., 1995; Guang et al., 2005; Malim et al., 1989).

Mammalian polyadenylation

In addition to splicing, 3′ end formation is critical for proper gene expression. In eukaryotes, 3′ end processing consists of two steps. First, the nascent transcript is cotranscriptionally cleaved by a multiprotein complex called the cleavage and polyadenylation specificity factor (CPSF). Following cleavage, poly(A) polymerase (PAP) appends a poly(A) tail to the upstream cleavage product. In mammalian cells, cleavage and polyadenylation depends on the AAUAAA consensus sequence found in the nascent transcript. This sequence motif is bound by CPSF, which subsequently directs endonucleolytic cleavage 10-30 nt downstream. CPSF

remains bound to the RNA following cleavage, and recruits poly(A) polymerase through a direct physical interaction (Keller et al., 1991). By itself, PAP has very weak affinity for RNA, and is essentially inactive in in vitro polyadenylation assays (Wahle et al., 1991). The physical interaction between PAP and CPSF allows polyadenylation to occur in a weakly processive manner (i.e. the addition of several adenosine residues followed by dissociation and rebinding). Once the nascent oligo(A) tail reaches a length of 10-14 nucleotides it is bound by the nuclear poly(A) binding protein (PABPN1). Like CPSF, PABPN1 forms a physical interaction with PAP (Wahle, 1991). When both factors are present, PAP is tightly tethered to the RNA, and polyadenylation is highly processive. Once the tail reaches a length of 200-300 nt, the interaction between PAP and CPSF is lost, and polyadenylation becomes weakly processive and dependent solely upon PABPN1 (Kuhn et al., 2009; Wahle, 1995).

For most RNAs, this marks the end of 3′ end processing. The mRNA is quickly exported to the cytoplasm, and PABPN1 is recycled back to the nucleus. However, some polyadenylated transcripts are retained in the nucleus. These include the predominantly nuclear long non-coding RNAs (lncRNAs) as well as inefficiently exported mRNAs. Because these transcripts presumably remain bound to PABPN1, they may continue to be slowly polyadenylated. We have termed this secondary polyadenylation event **“**hyperadenylation”**.**

Dual roles for polyadenylation in stability and decay

Polyadenylation is found in organisms from all three domains of life, and was presumably present in the last universal common ancestor 3.5 billion years ago. Despite this

shared ancestry, polyadenylation has evolved drastically different roles in eukaryotes compared to bacteria and archaea. Several differences in particular are worth mentioning.

In eukaryotes, the poly(A) tail is critically important for both translation and cytoplasmic stability. Once exported to the cytoplasm, the nuclear poly(A) binding protein is replaced by its cytoplasmic counterpart PABPC1 (Pab1 in *S. cerevisae*). PABPC1 stimulates translation by directly interacting with the translation initiation factor eIF4G at the 5′ end of the transcript. This stabilizes its binding to the transcript, which in turn promotes ribosome recruitment (Kahvejian et al., 2005). PABPC1 also enhances RNA stability by inhibiting deadenylation, the rate limiting step in cytoplasmic RNA decay.

In contrast to its well-known role in eukaryotic cytoplasmic RNA stability, polyadenylation in bacteria functions as a decay signal. Poly(A) tails in bacteria tend to be short (10-40nt), and are found on both full length transcripts and decay intermediates. Polyadenylation stimulates the activity of bacterial exoribonucleases such as RNase II and RNase R, which require 10-11 nt of single stranded, unstructured RNA in order to bind and initiate decay. Short poly(A) tails provide an unstructured landing pad that allows these enzymes to overcome local RNA secondary structure. Remarkably, this function remains conserved in mitochondria and chloroplasts (Slomovic and Schuster, 2011), which originated from bacterial ancestors more than one billion years ago.

Exosome

In eukaryotes, a multi-protein complex called the exosome supplies the majority of the RNA decay activity in the cell. The exosome is required for the processing and turnover of a

wide variety of substrates, including rRNA, tRNAs, snoRNAs, snRNA, mRNAs, and non-coding RNAs arising from pervasive transcription (Schneider and Tollervey, 2013). Originally discovered in yeast (Mitchell et al., 1997), the exosome is conserved throughout Eukarya and Archaea, and is distantly related to the bacterial RNase complex PNPase. The eukaryotic exosome consists of a central hexameric ring comprised of Rrp41, Rrp42, Mtr3, Rrp43, Rrp46, and Rrp45. These proteins share homology with bacterial RNase PH, but appear to have lost catalytic activity. Layered on top of the hexameric core are three RNA binding proteins: Rrp4, Rrp40, and Csl4. These 9 subunits are arranged in a barrel surrounding a central channel 8-10 Å wide, large enough to accommodate single-stranded, but not double-stranded RNA. Collectively, this 9-subunit complex is referred to as the exosome core or Exo-9, and bears close structural resemblance to the archaeal exosome. Exo-9 is catalytically inactive and RNase activity is supplied by one of several cofactors. In humans, the nuclear exosome can associate with both Dis3/Rrp44 and Rrp6. Dis3 is positioned at the base of the barrel and RNA must be threaded through the central channel to reach the active site. This requires ~30nt of naked, single-stranded RNA (Makino et al., 2013; Wasmuth and Lima, 2012). Rrp6 is positioned near the top of the barrel, but substrates must still be passaged through the central channel to access the catalytic site (Wasmuth et al., 2014).

These requirements impose considerable constraints on exosome-mediated decay. Consequently, the exosome requires the assistance of cofactors to facilitate decay. These cofactors perform two critical functions. First, they serve to reduce local secondary structure, either through helicase activity or through the addition of unstructured poly(A) tails. Second, cofactors can function as adaptors to increase the affinity of the exosome for various RNA substrates. Examples of the former include the TRAMP complex in budding yeast, which uses

both helicase and polyadenylation activity to stimulate the exosome. An example of the latter is Pab2-mediated decay in fission yeast. Pab2 binds poly(A) tails in the nucleus, and appears to stimulate RNA decay by increasing the affinity of the exosome for the target RNA. Both pathways are discussed below in detail.

TRAMP complex

Polyadenylation-mediated decay was initially thought to be unique to prokaryotes and organelles, but this changed with the discovery that poly(A)-tailed processing and decay intermediates accumulated in the nucleus of exosome mutants (Allmang et al., 1999; van Hoof et al., 2000). Subsequently, the putative poly(A) polymerase Trf4 was discovered in a genetic screen for genes involved in the degradation of hypomethylated $tRNAi^{Met}$ (Kadaba et al., 2004). Mutations in the methyltransferase Trm6 result in hypomethylated, and thus misfolded, $tRNAi^{Met}$ which is rapidly degraded by RNA quality control pathways. A screen for genetic suppressors identified Rrp44, an exosome component, and Trf4. Mutations in these factors stabilized hypomethylated $tRNAi^{Met}$ and rescued the slow growth phenotype in trm6 mutants. Importantly, $tRNAi^{Met}$ accumulated as a polyadenylated species in an exosome deletion strain, suggesting that $tRNAi^{Met}$ was polyadenylated by Trf4 as the initial step in RNA decay.

Subsequently, several groups discovered that Trf4 was a component of what was termed the TRAMP (Trf4/5-Air1/2-Mtr4p Polyadenylation) complex. TRAMP's polyadenylation activity can be supplied by either Trf4 or its closely related homolog Trf5. Trf4/5 cannot bind RNA directly, and thus polyadenylation is dependent on the RNA binding proteins Air1 or Air2. Mtr4, the third component of the TRAMP complex, is a 3′ to 5′ RNA helicase which functions to

relax local secondary structure. Following the discovery of the remaining members of the TRAMP complex, numerous other processing and decay substrates were identified. TRAMP has been shown to play an essential role in the processing and turnover of nearly all nuclear exosome targets in *S. cerevisae* (Callahan and Butler, 2010).

Initially, TRAMP was assumed to function analogously to the bacterial system: short poly(A) tails would serve as an unstructured platform for exonucleases, in this case the exosome, to bind and degrade the mRNA (Kadaba et al., 2004) (Vanacova et al., 2005). While this model is appealing, polyadenylation-assisted decay poses a problem for eukaryotes. How does the cell distinguish stabilizing poly(A) tails generated by canonical polyadenylation, from destabilizing poly(A) tails generated by TRAMP? Several explanations are possible. First, the answer may lie in differences in processivity between Trf4/5 and canonical Pap1. Trf4 is a distributive enzyme, and repeated cycles of binding and dissociation may allow the exosome a short window of opportunity to bind the naked 3′ end. In contrast, Pap1 is highly processive and remains bound throughout the polyadenylation process, protecting the RNA from the exosome. A second key difference between TRAMP and canonical polyadenylation is the extent of polyadenylation. Pap1 generates tails 60-80 nt in length, and these tails are coated with the poly(A) binding proteins Nab2 and Pab1. These proteins restrict further polyadenylation (Viphakone et al., 2008), and sterically hinder exosome-mediated decay. In contrast, TRAMP generates short poly(A) tails, 3-5nt in length. This is too short to bind either Pab1 or Nab2, which require 12nt and 20nt of poly(A), respectively (Sachs et al., 1987) (Viphakone et al., 2008). Finally, termination of canonical polyadenylation is coupled to nuclear export, whereas TRAMP substrates remain in the nucleus and susceptible to the exosome.

Despite the high conservation of Trf4/5's catalytic domain, it remains unclear if polyadenylation is even necessary for TRAMP-mediated decay. Mtr4 inhibits Trf4 activity once the tail reaches 3-5nt (Jia et al., 2011), consistent with the average tail length observed in vivo of 4-5 nt (Wlotzka et al., 2011). This falls well short of the ~30nt of unstructured, single-stranded RNA required for exosome-mediated decay. In vitro, TRAMP is able to stimulate $tRNAi^{Met}$ degradation in the absence of ATP (LaCava et al., 2005), suggesting that polyadenylation is dispensable. However, mutations in the catalytic site abolish activity towards the same substrate (Vanacova et al., 2005). Perhaps most surprisingly, a catalytic site mutant in Trf4 is able to rescue the levels of ~90% of decay targets in a trf4Δ background (San Paolo et al., 2009). These seemingly confusing results could be explained, at least in part, by the presence of Mtr4. Mtr4 possesses 3′ to 5′ helicase activity, and unwinding of local secondary structure may be sufficient, even in the absence of polyadenylation, to generate the ~30 nt of single stranded RNA needed by the exosome. Alternatively, perhaps only a subset of TRAMP targets requires active polyadenylation. Further research is needed to determine the precise role of polyadenylation in TRAMP-mediated RNA decay.

PABP-mediated decay

Recent work in fission yeast (*S. pombe*) has identified Pab2, a nuclear poly(A) binding protein, as a novel exosome cofactor. Pab2 has been implicated in snoRNA processing (Lemay et al., 2010), degradation of meiotic mRNAs during vegetative growth (Yamanaka et al., 2010), and turnover of nuclear pre-mRNAs (Lemieux et al., 2011). However, the precise mechanism by which Pab2 stimulates exosome activity remains somewhat unclear. Pab2 shares 47% identity

with its human homolog PABPN1, but unlike PABPN1, Pab2 is unable to stimulate polyadenylation. In fact, deletion of Pab2 leads to a global hyperadenylation phenotype (Perreault et al., 2007), suggesting that Pab2, like its *S. cerevisae* counterparts Nab2 and Pab1, normally suppresses polyadenylation. How then does Pab2 stimulate RNA decay? The most likely explanation is that Pab2 directly recruits the exosome to the substrate. Transient dissociation of Pab2 may then expose the poly(A) tail, allowing the transcript to be degraded. Consistent with this hypothesis, Pab2 co-immunoprecipitates with components of the exosome complex (Lemay et al., 2010).

Using PAN RNA as a model transcript to characterize mammalian nuclear RNA decay pathways

PAN (Polyadenylated nuclear) RNA is a 1.1kb transcript produced by Kaposi's sarcoma associated herpesvirus (KSHV). As its name suggests, PAN RNA is polyadenylated and mainly nuclear. PAN RNA is intronless, and like other unspliced transcripts which are poorly exported, should be rapidly degraded within the nucleus. Nevertheless, PAN accumulates to extremely high levels, comprising as much as 70% of the total polyadenylated RNA in the cell during the lytic phase of the virus. The high levels seen during lytic phase are dependent upon a 79 nt element just upstream of the poly(A) tail. This element was originally thought to regulate 3′ end formation and nuclear retention, and was named the expression and nuclear retention element (ENE) (Conrad and Steitz, 2005). However, subsequent work revealed that the ENE works by inhibiting nuclear decay, and this activity is dependent on a poly(A) tail. (Conrad et al., 2006). Importantly, the ENE is sufficient to stabilize intronless β-globin RNA in vivo (Conrad et al.,

2006). The ENE is predicted to form a stem loop with an internal U-rich bulge. The uridines within the central bulge were hypothesized to hybridize with the poly(A) tail, thereby protecting it from exonucleases (Conrad et al., 2006; Conrad et al., 2007). Subsequently, the crystal structure was solved of the ENE in complex with an A9 oligomer. Surprisingly, the ENE:poly(A) interaction forms a triple helix (Mitton-Fry et al., 2010). Subsequently, other studies found ENE-like structures in viral RNAs and two human RNAs (Brown et al., 2012; Tycowski et al., 2012; Wilusz et al., 2012), suggesting that poly(A) tail-sequestration may be a general mechanism to inhibit nuclear RNA decay pathways.

Importantly, the factors involved in PAN RNA decay, and mammalian nuclear RNA decay generally, remain unknown. In yeast, the major nuclear decay pathway involves TRAMP-mediated polyadenylation and the exosome. These factors are conserved in human cells, but their role in nuclear RNA decay is still unclear. Two observations regarding PAN suggest that, as in yeast, polyadenylation may play a role in nuclear decay. First, deletion of the ENE results in a less stable RNA with a longer poly(A) tail (Conrad et al., 2006). Second, the ENE physically interacts with the poly(A) tail and sequesters it within a triple helix structure (Conrad et al., 2006; Mitton-Fry et al., 2010). Importantly, the ENE is sufficient to stabilize intronless β-globin RNA in cis (Conrad et al., 2006), suggesting that the ENE inhibits a decay pathway which targets intronless transcripts generally. In this study, we investigate the factors and mechanisms underlying PAN and intronless β-globin decay.

Chapter 2: PABPN1 promotes RNA hyperadenylation and decay

Abstract

Control of nuclear RNA stability is essential for proper gene expression, but the mechanisms governing RNA degradation in mammalian nuclei are poorly defined. In this study, we uncover a mammalian RNA decay pathway that depends on the nuclear poly(A)-binding protein (PABPN1), the poly(A) polymerases (PAPs), PAPα and PAPγ, and the exosome subunits RRP6 and DIS3. Using a targeted knockdown approach and nuclear RNA reporters, we show that PABPN1 and PAPα, redundantly with PAPγ, generate hyperadenylated decay substrates that are recognized by the exosome and degraded. Poly(A) tail extension appears to be necessary for decay, as cordycepin treatment or point mutations in the PAP-stimulating domain of PABPN1 leads to the accumulation of stable transcripts with shorter poly(A) tails than controls. Mechanistically, these data suggest that PABPN1-dependent promotion of PAP activity can stimulate nuclear RNA decay. Importantly, efficiently exported RNAs are unaffected by this decay pathway, supporting an mRNA quality control function for this pathway. Finally, analyses of both bulk poly(A) tails and specific endogenous transcripts reveals that a subset of nuclear RNAs are hyperadenylated in a PABPN1-dependent fashion, and this hyperadenylation can be either uncoupled or coupled with decay. Our results highlight a complex relationship between PABPN1, PAPα/γ, and nuclear RNA decay and we suggest that these activities may play broader roles in the regulation of human gene expression.

Introduction

Prior to their export to the cytoplasm, nuclear pre-mRNAs must be capped, spliced, polyadenylated, and assembled into export-competent messenger ribonucleoprotein particles (mRNPs). Mistakes in any of these processes lead to aberrant mRNAs that may code for proteins with deleterious effects. As a result, cells have developed RNA surveillance or quality control (QC) mechanisms that preferentially degrade misprocessed transcripts (Doma and Parker, 2007; Fasken and Corbett, 2009; Schmid and Jensen, 2008b). While the mechanisms and factors involved in nuclear RNA quality control have been extensively studied in yeast models, these pathways remain largely uncharacterized in metazoans.

The addition of a poly(A) tail is essential for normal mRNA biogenesis, but polyadenylation can stimulate RNA QC pathways in *S. cerevisiae*. The yeast Trf4-Air2-Mtr4 polyadenylation (TRAMP) complex is a co-factor for the degradation of aberrant rRNA, tRNA, snoRNAs, snRNAs, long noncoding RNAs (lncRNAs), and mRNAs (LaCava et al., 2005; Schmid and Jensen, 2008b; Schmidt and Norbury, 2010; Slomovic and Schuster, 2011; Wyers et al., 2005). Decay is carried out by the nuclear exosome, a nine-subunit complex that associates with the Rrp6 and Dis3 nucleases (Schmid and Jensen, 2008a; Slomovic and Schuster, 2011). In addition to polyadenylation by TRAMP, a noncanonical poly(A) polymerase not involved in 3´end formation of mRNAs, predicted RNA QC targets can be hyperadenylated by Pap1, the canonical poly(A) polymerase (PAP) that polyadenylates mRNA 3´-ends (Grzechnik and Kufel, 2008; Hilleren et al., 2001; Qu et al., 2009; Rougemaille et al., 2007; Saguez et al., 2008), but whether this is directly linked to decay is unclear. Although the mechanisms linking polyadenylation with decay remain unknown, the nuclear poly(A)-binding protein (PABP),

Nab2, recruits the exosome to polyadenylated RNAs, suggesting that yeast PABPs couple hyperadenylation with decay (Schmid et al., 2012).

Few studies on mammalian nuclear RNA QC systems have been published, but some roles of the poly(A) tail in *S. cerevisiae* appear to be conserved in mammals. For example, mammalian TRAMP homologs promote polyadenylation and decay of aberrant rRNA and unstable promoter-associated transcripts (Lubas et al., 2011; Shcherbik et al., 2010; Slomovic et al., 2006). Furthermore, polyadenylation induced by a Kaposi's sarcoma-associated herpesvirus (KSHV) host shut-off protein results in the hyperadenylation and destabilization of host transcripts (Lee and Glaunsinger, 2009). Both yeast and mammalian mRNAs are hyperadenylated upon inhibition of bulk mRNA export (Hilleren and Parker, 2001; Jensen et al., 2001; Qu et al., 2009). In addition, knockdown of exosome components leads to the accumulation of oligoadenyated nuclear RNAs (West et al., 2006). Thus, certain aspects of poly(A) tail functions in nuclear RNA QC appear to be conserved in mammals, but little empirical evidence has been reported and mechanistic details remain largely unknown.

Our previous studies using the noncoding KSHV polyadenylated nuclear PAN RNA further support the idea that the poly(A) tail plays an important role in mammalian nuclear RNA decay. PAN RNA is a polyadenylated, capped, RNA polymerase II (pol II) transcript that accumulates to high levels in the nucleus, thereby making it a useful model to study nuclear RNA decay pathways. The high nuclear levels of PAN RNA depend on the presence of a 79-nt stability element near its 3′ end termed the ENE (Conrad et al., 2006; Conrad et al., 2007; Conrad and Steitz, 2005). The ENE interacts with the poly(A) tail in cis, protecting the transcript from degradation (Conrad et al., 2006; Conrad et al., 2007; Mitton-Fry et al., 2010). In addition, the ENE stabilizes an intronless β-globin reporter mRNA leading to its accumulation in the

nucleus. Due to the coupling of mRNA export with splicing, this intronless mRNA is less efficiently exported than its spliced counterparts (Cheng et al., 2006; Eckmann et al., 2011; Valencia et al., 2008) and therefore likely succumbs to the RNA QC machinery. These observations led to the model that inefficiently processed or exported transcripts are rapidly degraded in the nucleus, and implied that the presence of a poly(A) tail is linked to their degradation. Importantly, the factors involved in this decay pathway were not previously described.

Comparison of nuclear PABP functions in mRNA 3´-end formation and nuclear RNA decay suggest important mechanistic distinctions among fission yeast, budding yeast, and mammals. The human nuclear PABP, PABPN1, stimulates PAP processivity and controls poly(A) length in 3´ processing reactions in vitro (Eckmann et al., 2011). Recent studies suggest that PABPN1 also regulates alternative polyadenylation of specific mRNAs (de Klerk et al., 2012; Jenal et al., 2012). The *S. pombe* nuclear PABP, Pab2, is homologous to PABPN1, but is not related to *S. cerevisiae* Nab2. Pab2 promotes the exosome-mediated nuclear RNA decay of specific pre-snoRNAs, pre-mRNAs, and meiotic mRNAs (Chen et al., 2011; Lemay et al., 2010; Lemieux et al., 2011; Yamanaka et al., 2010). However, *S. pombe* Pab2 does not stimulate PAP activity (Eckmann et al., 2011) and deletion mutants in Pab2 are viable (Perreault et al., 2007), demonstrating that it is not essential for the polyadenylation of most mRNAs. Moreover, Pab2 or Nab2 depletion causes bulk poly(A) tail hyperadenylation (Hector et al., 2002; Perreault et al., 2007), whereas depletion of PABPN1 homologs in mouse and *Drosophila* decrease steady-state poly(A) tail lengths (Apponi et al., 2010; Benoit et al., 2005). These important distinctions in PABP function highlight the difficulties in extrapolating from one system to another, and underscore the need for detailed mechanistic studies in higher eukaryotes.

Here we define a pathway that promotes the degradation of polyadenylated nuclear RNAs in mammalian cells. This pathway depends on PABPN1, the canonical PAPs, PAPα and PAPγ, and the nuclear exosome components RRP6 and DIS3. The pathway targets ENE-lacking PAN RNA and intronless β-globin mRNA reporters, but does not degrade a spliced β-globin mRNA. Interestingly, PANΔENE RNA and intronless β-globin RNAs are polyadenylated upon PABPN1 knockdown or PAPα and PAPγ co-depletion, but the poly(A) tails are substantially shorter. In contrast, co-depletion of RRP6 and DIS3 stabilizes hyperadenylated forms of both intronless β-globin mRNA and PANΔENE RNA reporters and this hyperadenylation depends on PABPN1 and canonical PAP activity. Bulk RNA analysis reveals that a significant fraction of newly synthesized cellular polyadenylated RNAs is subject to PABPN1-dependent hyperadenylation. Efficiently spliced endogenous mRNAs appear to be unaffected by this decay pathway, but an endogenous polyadenylated nuclear noncoding RNA is, consistent with the interpretation that this pathway targets non-exported RNAs. Together, our results suggest that PABPN1 and the PAPs, PAPα and PAPγ, promote the exosome-mediated degradation of nuclear polyadenylated transcripts in human cells.

Results

PABPN1 is required for rapid PANΔENE RNA decay

To measure PANΔENE RNA stability, we used a well-characterized transcription pulse-chase assay (Conrad et al., 2006; Loflin et al., 1999; Sahin et al., 2010). In this assay, a plasmid that expresses PANΔENE under control of a tetracycline-responsive promoter (TetRP)(Figure 1A) is transfected into cells stably expressing a tetracycline-responsive transcriptional activator

in the presence of doxycycline (dox) to repress transcription. Transcription is induced for 2 hr by dox removal, then repressed by the readdition of dox, cells are collected over time and PANΔENE RNA levels are followed by northern blot. We previously demonstrated that PANΔENE RNA decay follows two-component decay kinetics (Conrad et al., 2006). A subpopulation of transcripts is degraded rapidly (defined herein as $t_{1/2} \leq 15$ min), while the remaining RNA pool is degraded more slowly. Regression analysis allows us to estimate the percentage of transcripts in each pathway and their corresponding half-lives. The presence of the ENE protects polyadenylated RNAs from the rapid decay pathway, but the cellular factors driving rapid decay have not been previously elucidated.

Interactions between the ENE and the poly(A) tail stabilize PAN RNA, suggesting that a poly(A) tail may promote rapid RNA decay. Because PABPN1 binds to poly(A) tails in the nucleus and the *S. pombe* PABPN1 homolog promotes RNA decay, we tested the effects of PABPN1 depletion on rapid PANΔENE decay. We used siRNAs to efficiently (~95%) knockdown PABPN1 in 293A-TOA cells (Sahin et al., 2010) (Figure 1B, Figure 2A and 2B). Transcription pulse-chase assays demonstrated that RNA decay was impaired upon PABPN1 depletion (Figure 1C). Two different PABPN1 siRNAs yielded similar results, consistent with a PABPN1-specific effect (Figure 1E, Figure 2C). Fitting the data to two-component exponential decay curves (Figure 1D) showed that nearly all of the RNA degraded rapidly in the control cells (~95%, $t_{1/2}$ ~7 min, Table 1), but predicted that only ~6.6% of the transcripts had $t_{1/2} \leq 15$ min in PABPN1-depleted cells. Thus, we conclude that PABPN1 is necessary for the rapid decay of PANΔENE RNA in vivo.

Because hyperadenylation often correlates with decay and PABPN1 stimulates PAP activity, we were interested in the poly(A) tail lengths of PANΔENE RNA in the presence or

absence of PABPN1. RNA samples taken immediately after the 2-hr transcription pulse (T=0) were cleaved with RNase H targeted by DNA oligonucleotide complementary to a region near the PANΔENE RNA 3′ end (NC30, Figure 1A), and the 3′ fragment was detected by northern blot using a 3′-end specific probe. A very broad smear was observed in the control cells (Figure 1F, lane 1) and this smear collapsed into a single discrete product upon addition of oligo dT_{40} to the RNase H reaction (Figure 1F, lane 3) demonstrating that the smear was due to heterogeneous poly(A) tail lengths. We estimated that the poly(A) tail distribution in the control cells corresponds to ~50-400 adenosines. Notably, this broad distribution of transcripts was lost when PABPN1 was depleted (compare lanes 1 and 2), but the transcripts maintained a sizable poly(A) tail of ~50-150 nt. Taken together, these data suggest that PABPN1 is necessary for the hyperadenylation and rapid decay of PANΔENE RNA.

PABPN1 is required for the rapid degradation of an intronless β-globin mRNA

To determine whether human mRNAs are subject to PABPN1-dependent decay, we initially examined the effects of PABPN1 knockdown on two TetRP-driven β-globin reporter constructs. One contains no introns (βΔ1,2) while the other retains the second intron of β-globin (βΔ1)(Figure 1A). Because pre-mRNA splicing promotes nuclear export, the inefficiently exported βΔ1,2 mRNA is unstable, presumably due to nuclear RNA QC (Conrad et al., 2006; Lei et al., 2011; Valencia et al., 2008). However, a fraction of the intronless β-globin is exported and this fraction is subject to slower decay in the cytoplasm (Conrad et al., 2006; Conrad and Steitz, 2005). Similar to PANΔENE, PABPN1 knockdown stabilized the intronless β-globin transcript (Figure 3A, 3B). Furthermore, the decay kinetics were consistent with two component exponential decay. In this case, ~73% of the βΔ1,2 was degraded rapidly in the control cells, but

only ~51% was subject to rapid decay upon PABPN1 knockdown (Table 1). In contrast, the spliced β-globin reporter mRNA was largely unaffected by PABPN1 knockdown and was stable over two-hour (Figure 3C) or eight-hour time courses (Figure 4). Thus, β-globin mRNA generated from an intronless gene is subject to PABPN1-dependent degradation, but the same mRNA produced from a spliced intron-containing gene is not.

We further examined the effects of PABPN1 knockdown on the poly(A) tail lengths of the intronless and spliced β-globin RNAs after a two-hour transcription pulse (Figure 3D). The intronless β-globin poly(A) tails closely resembled those of PANΔENE. That is, in the control cells, the poly(A) tails were longer and more heterogeneous than in the PABPN1 depleted cells (lanes 3 and 4). In contrast, the spliced β-globin mRNAs were less heterogeneous in either the presence or absence of PABPN1 and the difference in mobility between samples was smaller (lanes 1 and 2). Interestingly, the size of the intronless β-globin poly(A) tails in the absence of PABPN1 were more similar to those of the spliced β-globin RNA, consistent with the idea that the longer poly(A) tails were due to transcript hyperadenylation in control cells and not to hypoadenylation in the absence of PABPN1. Additionally, poly(A) tail analysis of the stable wild type version of PAN RNA revealed that it had a shorter poly(A) tail than PANΔENE, its unstable counterpart (data not shown). Taken together, these data suggest that unstable nuclear transcripts have longer poly(A) tails, and that PABPN1 is required for this poly(A) tail extension. For the purposes of this paper, we will describe the transcripts with PABPN1-dependent long poly(A) tails as "hyperadenylated". For PANΔENE, hyperadenylated transcripts have a poly(A) tail of roughly 150-400 nucleotides.

Polyadenylation by PAPα/γ promotes PANΔENE RNA decay

PABPN1 stimulates processive polyadenylation by PAPα (also known as PAP II, PAPOLA) (Eckmann et al., 2011) and stabilization of PANΔENE and βΔ1,2 by PABPN1 knockdown correlates with shorter poly(A) tails. Therefore, we reasoned that extension of poly(A) tails by PAPα may be linked to rapid PANΔENE decay. To assess the requirement of PAPα, and its close homolog PAPγ (neo-PAP, PAPOLG) (Kyriakopoulou et al., 2001; Topalian et al., 2001), for PANΔENE hyperadenylation and decay, we monitored PANΔENE stability upon siRNA-mediated knockdown of PAPα and PAPγ. We achieved robust knockdown of PAPγ, and substantial, though incomplete, knockdown of PAPα (Figure 5A). PANΔENE RNA was stabilized when both enzymes were depleted (Figure 5B, 5C), but not upon knockdown of either PAP individually (Figure 6A), suggesting functional redundancy between PAPα and PAPγ. In addition, different combinations of PAPα and PAPγ siRNAs also stabilized PANΔENE RNA, supporting the conclusion that stabilization was not due to off-target effects of siRNAs (Figure 6B). Notably, PAPα and PAPγ knockdown did not stabilize PANΔENE RNA to the same extent as PABPN1 knockdown. We suggest this is due to incomplete knockdown of PAPα compared to the highly efficient knockdown of PABPN1. Additionally, because PAPs are catalytic rather than stoichiometric factors, low protein levels may be sufficient for function. Even so, the protection of PANΔENE RNA from rapid decay upon PAPα and PAPγ co-depletion supports a role for the canonical PAPs in nuclear RNA decay.

We next examined the effects of PAPα and PAPγ depletion on PANΔENE poly(A) tail length (Figure 5D). Consistent with the stability data, knockdown of either PAP individually had little effect on RNA length (compare lanes 1-4), but co-depletion of PAPα and PAPγ reduced the poly(A) tail length of PANΔENE (lane 5). The decrease in poly(A) tail length was less than that observed upon PABPN1 knockdown (compare lanes 2 and 5), mirroring the effects on stability.

Importantly, no additive decreases in poly(A) tail length were observed when all three factors were depleted (lane 6), consistent with the proteins functioning in the same pathway. We conclude that PABPN1 and the canonical PAPs work in concert to hyperadenylate the unstable PANΔENE RNA.

The results above show that depletion of either PABPN1 or the canonical PAPs stabilizes transcripts that would otherwise be rapidly degraded. Because PABPN1 directly promotes PAP activity (Eckmann et al., 2011), our data suggest that PABPN1-dependent stimulation of hyperadenylation promotes decay of targeted transcripts. If this model is correct, overexpression of a mutant PABPN1 defective in PAP stimulation activity should block decay. To test this prediction, we introduced an L119A, L136A ("LALA") mutation into PABPN1 (Figure 5E). Previous work has shown that the resulting protein binds poly(A) with similar affinity to wild type protein, but is unable to stimulate polyadenylation by PAPα (Kerwitz et al., 2003; Kuhn et al., 2009; Kuhn et al., 2003). As predicted, PANΔENE RNA had a substantially shorter poly(A) tail upon LALA overexpression, similar to the sizes observed upon PABPN1 knockdown (Figure 6C, 6D). More importantly, the resulting RNA was more stable than the controls (Figure 5F, 5G), suggesting that PABPN1 must be able to stimulate polyadenylation in order to promote decay. Taken together, these results argue that PABPN1 stimulation of polyadenylation by PAP is important for its role in RNA decay.

Several pieces of data suggest that poly(A) tail extension precedes the rapid nuclear RNA decay observed here. First, the polyadenylation factors PABPN1 and the canonical PAPs are required for rapid decay. Second, the PAP-stimulating activity of PABPN1 is necessary for decay. Third, stabilization of PANΔENE or intronless β-globin RNA correlates with shorter poly(A) tails. Fourth, transient increases in PANΔENE length are sometimes observed at the

earliest time points after transcription inhibition (e.g. Figure 1C compare lanes 2 and 3). To test this hypothesis with another experimental approach, we examined the effects of the polyadenylation inhibitor cordycepin (3´-deoxyadenosine) on PANΔENE RNA. Addition of cordycepin coincident with the induction of transcription (T=-2) resulted in a robust stabilization of PANΔENE (Figure 5H, compare lanes 1-7 with 15-21, and Figure 5I). Comparison of PANΔENE RNA at time zero verified that cordycepin treatment resulted in shorter poly(A) tails (≤ 50 nt) than the control (Figure 6E). Because cordycepin acts through a chain termination mechanism, it can inhibit transcription as well as polyadenylation. The accumulation of transcripts after the two-hour pulse (Figure 5H, lane 16) and their short poly(A) tail lengths (Figure 6E) show that cordycepin primarily inhibited polyadenylation under our experimental parameters. Similar analyses demonstrated that intronless β-globin was stabilized by the addition of cordycepin (Figure 6F, 6G), whereas the stability of spliced β-globin was unaffected (Figures 6H, 6I). Thus, the generation of PANΔENE with short, cordycepin-terminated poly(A) tails increases PANΔENE half-life.

We next tested the effects of adding cordycepin simultaneously with transcription shutoff (T=0; Figure 5H, lanes 8-14, Figure 5I). Under these conditions, only polyadenylation occurring after transcription shut-off will be inhibited. Importantly, cordycepin must be converted to cordycepin triphosphate prior to termination of polyadenylation, so the efficiency of inhibition may be limited by the in vivo kinetics of substrate phosphorylation. Even so, we observed a protection of PANΔENE RNA from rapid decay (Figure 5H and 5I). Regression analysis suggests that only ~67% of transcripts undergo rapid decay when cordycepin is added at T=0 compared to ~90% in untreated samples (Table SI). These data suggest that poly(A) tail extension, and not strictly length, may be important for rapid RNA decay in vivo (see

Discussion). However, we acknowledge that a terminal 3´-deoxyadenosine could inhibit decay factors requiring a 3´ hydroxyl group or the presence of cordycepin in cells may have additional indirect consequences on cellular metabolism (e.g. (Holbein et al., 2009)).

The nuclear exosome is required for PANΔENE rapid decay

Given its roles in RNA QC pathways, we next tested whether the nuclear exosome is required for PABPN1-mediated decay. We used siRNAs to knockdown both catalytic cofactors of the nuclear exosome, RRP6 and DIS3 (Figure 7A). Using our transcription pulse-chase assay, we found that PANΔENE RNA was stabilized upon co-depletion of RRP6 and DIS3 (Figure 7B, 7C). Individual knockdowns had only marginal or no effect on PANΔENE stability (Figure 8A), consistent with prior reports of functional redundancy between RRP6 and DIS3 (Kiss and Andrulis, 2011; Preker et al., 2008; Tomecki and Dziembowski, 2010). In addition, different combinations of RRP6 and DIS3 siRNAs also stabilized PANΔENE RNA, diminishing concerns of off-target effects of siRNAs (Figure 8C).

If PABPN1 stimulates polyadenylation by PAPα or PAPγ as a prerequisite to exosome-mediated decay, then knockdown of RRP6 and DIS3 should lead to the accumulation of hyperadenylated (~150 to ~400nt) RNAs. Indeed, upon exosome depletion, the poly(A) tails were equivalent in length to those from siControl treated cells (Figure 7B and Figure 8B), rather than the shorter forms seen upon PABPN1 or PAPα/PAPγ depletion. Interestingly, upon overexposure, a subset of transcripts migrated as broad heterogeneous smears in the RRP6 and DIS3 depleted cells (Figure 7B). Following treatment with RNase H and oligo(dT), the low mobility smears disappeared (Figure 7D), verifying that the increase in size was due to excess polyadenylation. Because these extremely hyperadenylated (>400nt) forms were not seen in the

control samples, they are unlikely to be natural decay intermediates. Rather, these abnormally long tails may be an experimental artifact of exosome depletion due to the decoupling of polyadenylation and exosome-mediated degradation. Taken together, these data support a role for the exosome in rapid PANΔENE decay. In addition, they are consistent with the model that poly(A) tail extension precedes PANΔENE decay.

Rapid decay of intronless β-globin requires canonical PAP and exosome activity

To examine whether the canonical PAPs and exosome were required for mRNA QC, we co-depleted PAPα and PAPγ or RRP6 and DIS3 and monitored intronless β-globin RNA stability. Consistent with the PANΔENE results, intronless β-globin was stabilized upon depletion of these factors (Figure 9A, 9C). In addition, exosome depletion resulted in the accumulation of low mobility transcripts (Figure 9A, double asterisks) due to excessive hyperadenylation, as confirmed by RNaseH/dT analysis (Figure 9A, right panel). In contrast to the intronless reporter, spliced β-globin stability and poly(A) tail length were largely unchanged by either knockdown (Figure 9B). If PABPN1 and the canonical PAPs hyperadenylate RNAs to target them for exosome-mediated decay, then the hyperadenylation observed upon RRP6 and DIS3 co-depletion should be lost in the absence of PABPN1 or the PAPs. Indeed, exosome depletion no longer led to the accumulation of hyperadenylated or extremely long poly(A) tails when PAPs or PABPN1 were co-depleted (Figure 9D, compare lanes 13-16 and 17-20 to lanes 9-12; see Figure 10A for quantification). Similar results were observed for PANΔENE (Figure 10B). Taken together, these data support the conclusion that PABPN1 promotes PAPα or PAPγ-dependent hyperadenylation of nuclear polyadenylated RNAs leading to their degradation by RRP6 or DIS3.

PABPN1 is responsible for the hyperadenylation of newly made transcripts in vivo

We next evaluated the impact of PABPN1 on bulk cellular polyadenylated RNAs. To ensure that we monitored RNAs synthesized after functional knockdown of each factor and thereby diminish ambiguities that arise from long-lived transcripts generated prior to knockdown, we developed an in vivo RNA labeling technique to monitor newly made poly(A) tails (Figure 11A). The cells were treated for two hours with the modified nucleoside 5-ethynyluridine (EU), which is efficiently incorporated into transcripts by elongating RNA polymerases (Jao and Salic, 2008). After harvesting cells and RNA extraction, we biotinylated the labeled transcripts with "click" chemistry and purified labeled RNA on streptavidin (SA) beads. The selected RNAs were digested to completion with RNase T1, a G-specific endonuclease that digests total RNA, but leaves the poly(A) tails intact. The resulting poly(A) tails were detected by northern blot using a radiolabeled oligo(dT) probe. Importantly, no poly(A) tails were detected in the absence of EU treatment (Figure 11B, lanes 1 and 2), confirming the specificity of our purification. Cells treated with a control siRNA displayed a broad poly(A) tail distribution (~100-400 nt) (lane 3) whereas PABPN1 depletion led to the selective loss of the longest poly(A) tails (>~200 nt) (lane 4). These data mimic what we observed with PANΔENE and intronless β-globin transcripts in that hyperadenylation did not occur when PABPN1 was depleted. However, it remains formally possible that the RNAs with longer poly(A) tails disappeared due to their degradation. Similar results were observed with shorter EU pulses, although the degree of hyperadenylation in the control cells was slightly diminished (Figure 12A). When we collected nuclear and cytoplasmic fractions after the EU

pulse, the longer poly(A) tails were primarily observed in the nuclear fraction, as expected if the RNAs are hyperadenylated in a PABPN1-dependent manner (Figure `1B).

We also tested other factors involved in PANΔENE and intronless β-globin rapid decay in this assay. Similar loss of long poly(A) tails was observed when PAPα and PAPγ were co-depleted (Figure 11C, lanes 2 and 3). We additionally found that RRP6 and DIS3 co-depletion increased the accumulation of transcripts with hyperadenylated tails, (compare lanes 2 and 4), particularly those with very long poly(A) tails (>400 nt) (Figure 11D). Importantly, this hyperadenylation was lost if PAPα and PAPγ were co-depleted with the exosome subunits (Figure 11C, lanes 4 and 5, Figure 11D). This response of bulk poly(A) tails to PABPN1, PAP, and exosome depletion closely matches our observations with PANΔENE and intronless β-globin. These data suggest that PABPN1, PAPs, and the exosome are active in the hyperadenylation and decay of a subset of endogenous human transcripts.

To investigate the role of PABPN1 on specific endogenous RNAs, we first examined the effects of PABPN1 depletion on the levels of several mRNAs and nuclear lncRNAs. Upon depletion of PABPN1, the abundance of GAPDH, β-actin, or ARGLU1 mRNAs was unaffected (Figure 13A and 13B). In addition, the nucleocytoplasmic distribution of GAPDH or β-actin mRNAs was unaltered by PABPN1 knockdown (Figure 14A and 14B). These mRNAs were examined at steady state, so it is formally possible that pre-existing, stable mRNAs may mask a general effect of PABPN1 knockdown on mRNA synthesis or decay. To test this, we knocked down PABPN1 and then induced the expression of two interferon stimulated genes, OAS2 and viperin, by adding interferon-α to the cells for 5 hr. Consistent with the steady-state analysis, PABPN1 knockdown had no effect on the accumulation of either IFN-stimulated mRNA (Figure 13C).

Next, we examined the steady-state levels of three abundant and well-studied nuclear lncRNAs, MALAT1, NEAT1, and XIST. Each of these RNAs undergoes a slightly different maturation process. The MALAT1 RNA has no introns, and its 3′ end is produced by RNase P-mediated endonucleolytic cleavage and stabilized by formation of a triple helix structure (Brown et al., 2012; Wilusz et al., 2008; Wilusz et al., 2012). As a result, MALAT1 lacks a poly(A) tail, and thus should not be subject to PABPN1-mediated decay. NEAT1 is also intronless, but it possesses a conventional poly(A) tail. In this sense, NEAT1 is most analogous to our PANΔENE reporter RNA. Finally, XIST is polyadenylated by the canonical cleavage and polyadenylation machinery, but has multiple introns. PABPN1 depletion had distinct effects on the levels of each of these nuclear noncoding RNAs (Figure 13A and 13B). We saw no significant changes in MALAT1 or XIST lncRNA levels (Figure 13A and 13B). In marked contrast, NEAT1 levels increased ~5-fold when PABPN1 was depleted. These results are in agreement with a recent study showing that the steady state levels of most mammalian mRNAs are unaffected upon PABPN1 knockdown but that a subset of nuclear noncoding RNAs, including NEAT1, increased in abundance (Beaulieu et al., 2012).

We also examined the relative poly(A) tail lengths of several transcripts in the presence or absence of PABPN1. Knockdown of PABPN1 had no effects on the mobility of GAPDH mRNA, β-actin mRNA, or the MALAT1 lncRNA (Figure 13D). However, upon PABPN1 depletion, the mean poly(A) tail lengths of NEAT1 were shorter and less heterogeneous than in the control cells, mirroring our observations with PANΔENE and intronless β-globin. Interestingly, in the case of the XIST lncRNA, the mean poly(A) tail distribution was shorter in the absence of PABPN1 (Figure 13D), but the steady-state levels increased only marginally

(~30%) (Figure 13B). Thus, it appears that PABPN1-dependent nuclear hyperadenylation can be separated from the subsequent decay pathways in specific cases (see Discussion). These results are consistent with the reporter assays and support the conclusion that PABPN1 can promote the hyperadenylation and decay of endogenous nuclear transcripts.

Discussion

In this study, we used a viral nuclear RNA and intronless β-globin reporters to identify components of a rapid human nuclear RNA decay pathway. Through a targeted knockdown approach, we identified PABPN1, the canonical PAPs, PAPα and PAPγ, and the nuclear exosome components RRP6 and DIS3 as central players in this rapid decay pathway. We propose that this pathway promotes the decay of nuclear mRNAs undergoing RNA QC and polyadenylated nuclear noncoding RNAs.

Poly(A) tail extension and nuclear RNA decay

The relationship between RNA polyadenylation, hyperadenylation and RNA decay is complex. Hyperadenylated RNAs are often observed when RNA decay factors or other RNA processing factors are compromised, but it is difficult to empirically determine whether the hyperadenylation promotes RNA decay or if hyperadenylated transcripts appear as a consequence of manipulating nuclear RNA metabolism. Several of our observations suggest that poly(A) tail extension is mechanistically linked to decay. Knockdown of either PABPN1 or co-depletion of the PAPs leads to more stable RNAs with shorter poly(A) tails. Importantly, the tails observed upon PABPN1 knockdown are not completely lost, but are still ~50-150 nt for PANΔENE RNA (Figure 1F) and ~100-200 nt in the case of newly made bulk poly(A) RNA

(Figure 11). Overexpression of the PABPN1 LALA mutant, which binds RNA but cannot stimulate polyadenylation, strongly stabilized PANΔENE RNA and led to shorter poly(A) tails (Figure 5F, 5G). Thus, PABPN1-dependent extension of the poly(A) tails to lengths longer than ~200 nt may precede decay. Consistent with this, exosome depletion led to the stabilization of hyperadenylated transcripts, some with very long tails, and generation of hyperadenylated tails depends on PABPN1 and the PAPs (Figures 7 and 9). In addition, when examined shortly after transcription shut-off (t=15 min), we sometimes observed that PANΔENE RNA became slightly longer immediately prior to its destruction (for example, compare lanes 2 and 3 in Figure 1B) suggesting we were able to detect the poly(A) tail extension prior to transcript decay. Finally, inhibition of poly(A) tail extension by cordycepin inhibited transcript decay, even when cordycepin was added coincident with transcription shut-off (Figure 5H and 5I). Importantly, a terminal 3´deoxynucleotide has no effect on exosome activity in vitro (C. Lima, personal communication), supporting our interpretation that the stability effects seen here are due to the inhibition of polyadenylation, rather than exonucleolytic decay directly. Admittedly, we cannot completely rule out indirect effects of cordycepin on RNA metabolism. However, when taken together, our data are most consistent with the model that poly(A) tail extension is directly involved with transcript decay and RNA QC.

An evolutionarily conserved role for nuclear poly(A) binding proteins in RNA decay

The present work complements recent studies in both fission and budding yeast pointing to novel roles for nuclear PABPs in RNA decay. In budding yeast, the nuclear poly(A)-binding protein, Nab2, binds to hyperadenylated tails generated by TRAMP and recruits Rrp6 to degrade the transcript (Schmid et al., 2012). Additionally, Nab2 autoregulates the *NAB2* mRNA by

promoting Rrp6-mediated decay (Roth et al., 2009; Roth et al., 2005). However, Nab2 is not a homolog of PABPN1 and no *S. cerevisiae* PABPN1 homolog has been described (Winstall et al., 2000). Moreover, Nab2 does not stimulate yeast PAP activity in vitro and loss of Nab2 leads to hyperadenylation of bulk transcripts rather than mirroring the loss of long poly(A) tails observed with PABPN1 depletion (Figure 11) (Apponi et al., 2007; Dheur et al., 2005; Hector et al., 2002). Nab2 homologs have been identified in metazoans (Kelly et al., 2007; Pak et al., 2011), but their functions remain poorly understood. Therefore, it remains unclear whether Nab2 is an appropriate model for exploring PABPN1 function.

Perhaps a better model for human PABPN1 is the fission yeast Pab2 protein, which shares 47% identity and 66% similarity with PABPN1 (Perreault et al., 2007). Deletion of Pab2 stabilizes pre-snoRNAs, a subset of pre-mRNAs, and meiosis-specific mRNAs (Chen et al., 2011; Lemay et al., 2010; Lemieux et al., 2011; St-Andre et al., 2010; Yamanaka et al., 2010). Similar to our results, Pab2-mediated decay requires polyadenylated targets, the canonical PAP (Pla1), and Rrp6. However, there are important mechanistic distinctions between PABPN1 and Pab2 in decay. For example, while human PABPN1 stimulates PAP processivity, Pab2 does not stimulate Pla1 in vitro (Eckmann et al., 2011). Most strikingly, deletion of Pab2 or PABPN1 has opposing effects on poly(A) tail lengths. Deletion of Pab2 leads to increases in bulk poly(A) lengths and to the hyperadenylation of at least some of its targets (Chen et al., 2011; Lemay et al., 2010; Perreault et al., 2007), whereas PABPN1 depletion decreases bulk poly(A) tail lengths (Figure 11)(Apponi et al., 2010). Interestingly, the *Drosophila* PABPN1 homolog mimics the human PABPN1 in that it stimulates its cognate PAP and its depletion leads to poly(A) tail shortening (Benoit et al., 2005; Benoit et al., 1999; Juge et al., 2002), suggesting that these

functions may be restricted to metazoans. As described above, we favor a model in which PABPN1-mediated stimulation of polyadenylation constitutes an important part of a nuclear RNA decay pathway and PABPN1 may additionally recruit the exosome. In contrast, the work in *S. pombe* suggests the primary function of Pab2 in decay is to recruit the exosome to polyadenylated transcripts. Upon Pab2 deletion, the exosome is less efficiently recruited, resulting in increased Pla1 activity and hyperadenylation. Thus, while nuclear poly(A)-binding proteins from fission yeast, budding yeast, and humans all promote nuclear RNA decay, the mechanistic details appear to differ considerably.

Targets of PABPN1-mediated RNA decay

A recent RNA-seq study demonstrated that PABPN1 depletion had little effect on the levels of most transcripts, but a subset of nuclear lncRNAs were stabilized (Beaulieu et al., 2012). These results are entirely consistent with our data with PAN RNA, a viral lncRNA. In addition, we propose this pathway targets export-deficient mRNAs, similar to our intronless β-globin mRNA. These transcripts would remain undetected in RNA-seq studies because they will represent only a small subpopulation of any specific mRNA. Interestingly, a large fraction of newly made RNAs acquire long poly(A) tails in a PABPN1-dependent fashion (Figure 11B). If these transcripts all represent (pre-)mRNAs targeted for elimination by RNA QC pathways, then it is noteworthy that the cell generates so many aberrant RNAs. Moreover, it seems unlikely that these transcripts are primarily lncRNAs as these are generally low abundance transcripts. Alternatively, some of the newly made RNAs with long poly(A) tails may be hyperadenylated in a manner that is uncoupled from the subsequent decay processes. In fact, this appears to be the case for XIST, as we observed shorter poly(A) tails upon PABPN1 depletion, but the steady-state

levels increased only marginally (Figure 13B, 13D). Interestingly, PAN ENE, intronl globin and NEAT1 are all single-exon transcripts, whereas XIST has multiple exons, consistent with previous reports indicating a mechanistic link between splicing factors and RNA stability (Conrad et al., 2006; Hautbergue et al., 2009; Lei et al., 2011; Stubbs et al., 2012; Zhao and Hamilton, 2007). Alternatively, XIST and other stable polyadenylated nuclear RNAs may have evolved other cis-acting mechanisms to promote RNA stability. Indeed, wild-type PAN RNA uses the ENE for protection from nuclear decay factors. Future experimentation will focus on determining the identity and regulation of these abundant PABPN1-mediated hyperadenylated transcripts.

The role of PABPN1 in the initial polyadenylation event

PABPN1 and PAPα are generally thought to be required for the initial polyadenylation of nascent transcripts. However, under conditions in which we observed effects on hyperadenylation and RNA stability of our reporters and the endogenous NEAT1, we see little or no effect of PABPN1 knockdown on steady-state mRNA levels, mRNA nucleocytoplasmic distribution or in the accumulation of IFN-stimulated mRNAs. Moreover, changes in poly(A) tail lengths of spliced β-globin reporters or bulk newly-made cytoplasmic poly(A) tails were not as large as those observed for nuclear RNAs. There are several possible interpretations of our seemingly contradictory data. First, PABPN1 and PAP are required for both initial polyadenylation and subsequent hyperadenylation, but our knockdown was sufficient only to affect hyperadenylation. Because the PAPs are catalytic factors, low amounts may be sufficient for activity. Additionally, low levels of PABPN1 can stimulate PAPs in vitro (Wahle, 1995), so it is possible that the PABPN1 remaining after knockdown could provide function. Second, the

distributive PAPα activity observed in the absence of PABPN1 may be sufficient to produce the poly(A) tails observed over a two-hour transcription pulse. Importantly, our assays do not measure the rate of the initial polyadenylation reaction, so it is possible that there is a reduced rate of initial PAP activity in the absence of PABPN1, but it is beyond our experimental parameters to measure it. However, even though the poly(A) tails reach lengths of ~50-150 nt during the 2-hr transcription pulse, they are not extended during the subsequent 2-hr chase (Figure 1, lanes 9-14), suggesting that the shorter poly(A) tails are not simply the result of slower polyadenylation or that further extension is counterbalanced by deadenylases. Third, other proteins may substitute for PABPN1 in its absence. Indeed, a recent study found that when PABPN1 is depleted, the cytoplasmic poly(A) binding protein PABPC4 is relocalized to the nucleus where it associates with the polyadenylation machinery (Bhattacharjee and Bag, 2012). Fourth, PABPN1-mediated stimulation of PAP processivity may not be absolutely required for the formation of initial poly(A) tails in vivo. The evidence that PABPN1 participates in the initial cleavage and polyadenylation event is derived from in vitro experiments. These experiments show that PABPN1, synergistically with the cleavage and polyadenylation specificity factor (CPSF), promotes processive PAPα activity in the initial polyadenylation reaction (Eckmann et al., 2011). Once the tail reaches ~250 nt, PABPN1 alone is largely required for further polyadenylation (Kuhn et al., 2003; Wahle, 1995). Therefore, perhaps CPSF alone or with other unknown factors is sufficient to promote PAPα processivity during the primary mRNA 3´-end formation in vivo. Importantly, even if PABPN1 is not required for the initial polyadenylation reaction, our results are remarkably consistent with the in vitro evidence supporting a role for PABPN1 in promoting polyadenylation beyond ~250 nt. Distinguishing among these

possibilities will provide important insight into the complex roles of nuclear polyadenylation and poly(A)-binding proteins in mRNA biogenesis and decay.

Materials and methods

Plasmids and cell culture

The TetRP-driven PANΔENE and β-globin reporter constructs were described previously (TRP-PANΔ79)(Conrad et al., 2006; Conrad et al., 2007). The myc-PABPN1 overexpression construct was a gift from Dr. Lynne Maquat (University of Rochester)(Hosoda et al., 2006). Point mutations were generated by PCR using myc-PABPN1 as a template. PCR fragments were cloned into myc-PABPN1 to generate myc-tagged LALA. Positive clones were confirmed by sequencing. 293A-TOA cells (Sahin et al., 2010) were grown in Dulbecco's modified Eagle's medium (Sigma) containing 10% Tetracycline-free fetal bovine serum (Clontech), 1X penicillin/streptomycin (Sigma), 2 mM L-glutamate, and 100 µg/mL G418.

Transfections

Cells were transfected with 10 nM siRNA (Silencer Select, Ambion; Table 2) using RNAiMAX transfection reagent (Invitrogen) according to the manufacturer's instructions. One day following transfection, ~100% confluent cells were diluted to new plates to allow the cells to divide for an additional 24-48 hours, after which they were transfected with the appropriate reporter constructs. For DNA transfections, cells were transfected with TransIT-293 (Mirus) with the following modifications to the manufacturer's protocol. For transfection of a single well from a 12-well plate, 0.5 µL TransIT-293, 10 µL Opti-MEM reduced serum media (Gibco), and 0.2 µg of DNA were used. A typical transfection consisted of 0.1 µg pcDNA3 and 0.1 µg reporter construct. Twenty-four hours following transfection cells were harvested using TRI Reagent (Molecular Research Center).

RNase H analysis

Total RNA (1-3 μg) was incubated at 37˚C for 1 hour in a 20 μL mixture containing 0.3 U RNase H (Promega), 20 U RNasin (Promega), 300 ng poly(A) RNA (Sigma), 10 mM dithiothreitol, 100 mM KCl, 10 mM $MgCl_2$, and 20 mM Tris-HCl (pH 7.0), and plus or minus 0.5 μM dT_{40}. We found that the addition of poly(A) was necessary to counter a trace deadenylation activity present in the reaction. In Figure 1, internal cleavage of PANΔENE RNA was directed by the addition of 0.5 μM NC30 (5´ CAATGTTCTTACACGACTTTGAAACTTCTGACAAATGCC 3´). Following digestion, 180 μL of G50 buffer (0.25% SDS, 0.3 M NaOAc, 20 mM Tris pH 6.8, and 2 mM EDTA) was added to stop the reaction. After PCA extraction, the RNA was precipitated in 70% ethanol and run on a 1.6% agarose gel alongside a 0.1 to 2 kb RNA ladder (Invitrogen). Blots were probed with an end-labeled oligo probe specific to the 3´ end of PANΔENE RNA. Poly(A) tail lengths were estimated by comparison with the size marker. For analysis of endogenous transcripts 90 μg total RNA was treated with RNase H and DNA oligos complementary to the 3´ ends of MALAT1 (5´ AAGCACCGCTTGAGATTTGGG 3´), NEAT1 (5´ TTCCAAACTGATTTTAGGTGA 3´), and XIST (5´ TAGACAAACCTTGTAAATGC 3´). The final concentration of each DNA oligo was 0.5 μM in a 60 μL reaction. Following RNase H treatment, RNA was poly(A) selected according to standard protocols and half of each sample was treated with RNase H and oligo(dT) and the 3´ cleaved fragments were detected by northern blot.

Transcription pulse-chase assays

Transcription pulse-chase assays were performed as previously described (Conrad et al., 2007; Sahin et al., 2010) with the following exceptions. Twenty-four hours post-siRNA transfection,

cells were plated on 12-well plates. Twenty-four to 48 hours later, depending on the targeted mRNA, cells were transfected with the appropriate DNA constructs. Cells were grown in the presence of 5 ng/mL of dox (Sigma-Aldrich) to repress transcription. Twenty-four hours after DNA transfection, cells were washed twice with 1X Dulbecco's Phosphate Buffered Saline (PBS) with calcium/magnesium (Sigma-Aldrich), and cells were incubated in dox-free media for two hours. Dox was added to a final concentration of 50 ng/mL to repress transcription. For the experiments in Figure 5 and 6, 20 μg/mL cordycepin (Sigma-Aldrich) was added upon initiation of the two-hour pulse and remained in the media after dox addition. Total RNA was harvested at the appropriate times using TRI Reagent, and analyzed by northern blot. Membranes were probed with radiolabeled antisense transcripts for PAN RNA or β-globin, and subsequently for 7SK RNA.

Regression analysis

We fit transcription pulse-chase data to two-component exponential decay curves with Prism 5 software (GraphPad) using default parameters with two additional constraints. First, the decay constant for the rapid decay (K_{fast}) pathway was set to ≥ 0.046 corresponding to a $t_{1/2} \leq 15$ (Conrad et al., 2006). Second, we set the plateau value to zero. While all the replicates were used in regression analysis ($n \geq 3$), the reported R^2 values (Table SI) are derived from the fit of the mean values only. To focus on the rapid decay process, we took samples for up to 2hrs. Therefore, we report $t_{1/2}$ values for the slow decay pathway as > 120 min if they exceed the sampling time. All decay curves were fit using this method, except for the siRRP6/siDIS3 data in Figure 9C and LALA overexpression data in Figure 5G which did not conform to this model and

Figures 6G and 6I which had no data points for times <1hr thereby ruling out the definition of rapid decay pathway.

Immunoblotting

Protein was harvested using TRI Reagent according to the manufacturer's instructions. Protein was separated on SDS-PAGE and western blotted using standard procedures. The following antibodies were used: anti-PABPN1 (Abcam; ab75855), anti-PAPα (Abcam; ab126934), anti-PAPγ (Novus Biologicals; NBP1-30061), anti-RRP6 (EXOSC10) (Abcam; ab50558), anti-DIS3 (Bethyl Laboratories; A303-765A), anti-β-actin (Abcam; ab6276). Quantitative westerns were performed using infrared detection with an Odyssey Fc and quantification was performed using ImageStudio software (LI-COR Biosciences).

Metabolic labeling and selection of newly made RNAs

Click-iT Nascent RNA Capture Kit (Invitrogen) was used with modifications described in [52]. Briefly, two to four days following knockdown, cells were exposed to 200 μM EU for two hours. RNA was harvested using TRI Reagent. EU-containing transcripts were selected from 500 ng of total RNA using the biotinylation and SA selection described in Grammel et al. (2012). For the experiment shown in Figure 12A, three different pulse times (120', 60', and 30') were used. Because longer pulse times would be expected to label more RNA, we increased the amount of RNA used from the shorter pulses (0.5 μg for 120 min, 1 μg for 60 min, and 2 μg for 30 min). For the experiments in Figure 12, a biotinylated DNA oligonucleotide (5´ GATATTGAATCGAAAATCATATCTTTGATAATAGACTACTCAAGACTTTGTCCCGAT TCTCCTTTAAACTTGAAG-Btn 3´) was added to the reaction in order to control for

differences in recovery. The loading control was detected with a complementary DNA oligo (5´ GTCTTGAGTAGTCTATTATCAAAGATATGATTTTCGATTC 3´).

Bulk poly(A) RNA analysis

Total or SA-captured biotinylated RNA was incubated at 37˚C for 15 minutes in a 30 µL reaction containing 20 mM TrisHCl (pH 6.8), 8 U RNasin and 150 U RNase T1 (Ambion). The reaction was stopped by adding 170 µL G50 buffer containing 0.1 mg/mL Proteinase K (Fisher) and incubated at 37˚C for 30 minutes. After PCA extraction and ethanol precipitation, products were separated on a 1.8% agarose-formaldehyde gel, and analyzed by northern blot using a dT_{40} probe end-labeled with T4 polynucleotide kinase.

Northern blotting

Standard northern blotting techniques (Church and Gilbert, 1984) were used to probe for PAN, β-globin, and 7SK. For the experiments performed in Figure 13, we made several modifications. In order to detect multiple RNAs per gel, we made cleavable riboprobes that could be stripped from the membrane. To do this, we generated probes incorporating a partial phosphorothioate backbone, which can be subsequently cleaved by treatment with iodine. Upon cleavage, the full-length probe is reduced to a number of smaller fragments, which can be easily removed with washing. A typical probe reaction consisted of 40 mM Tris (pH=7.5) 6 mM $MgCl_2$, 4 mM spermidine, 10 mM DTT, 200 ng T7-driven template, 0.5 mM ATP, CTP, and GTP, 2 U/µL RNasin, 50 µM UTPαS (TriLink Biotechnologies), 25 µCi of α-^{32}P-UTP (800 Ci/mmol), and T7 RNA polymerase. After visualizing the target RNA, the membrane was incubated at 65°C for 30 minutes in a solution of 200 mM sodium phosphate buffer (pH=7.2), 50% formamide, 7%

SDS, and 6.25 mM I_2. Immediately following the cleavage reaction, the probe was stripped by washing the membrane twice in boiling 0.5% SDS.

Interferon and qPCR

Two days following siRNA transfection, cells were treated for five hours with 100 U/mL human interferon-α A (PBL Interferon Source). RNA was harvested, and cDNA was made with random hexamers as primers using standard molecular biology techniques. Quantitative real-time PCR was performed using iTaq Universal SYBR Green Supermix (BIO-RAD). Primers: GAPDH 5´ AGCCTCAAGATCATCAGCAATG, GAPDH 3´ ATGGACTGTGGTCATGAGTCCTT. Viperin (QT01005256) and OAS2 (QT00005271) primers were obtained from Qiagen.

Nucleocytoplasmic Fractionation

Two days following knockdown, cells from 10 cm plates were trypsinized and resuspended in ice-cold media. Cells were pelleted with gentle centrifugation at 500g for 3 minutes at 4° C, and then washed with ice-cold PBS. Following the wash step, cells were resuspended in 600 µL of ice cold Buffer I (0.32 M Sucrose, 3 mM $CaCl_2$, 2 mM $MgCl_2$, 0.1 mM EDTA, 10 mM Tris (pH=8.0), 1 mM DTT, 0.4 U/ µL RNasin, and 0.5% Igepal) and incubated for 5 minutes on ice. The cell lysates were spun at 500g for 5 minutes at 4° C. The supernatant ("cytoplasmic" fraction) was added to 5 mL TRI Reagent and the pellet ("nuclear" fraction) was resuspended in 600 µL Buffer I before addition of 5 mL TRI Reagent. In order to ensure consistent loading between samples, we poly(A) selected 40 µg of RNA from the cytoplasmic fraction and 10 µg of RNA from the nuclear fraction. To control for the quality of the fractionation, an aliquot of non-

poly(A) selected RNA was probed for pre-ribosomal RNA using a DNA oligo probe (5´ AACGATCAGAGTAGTGGTATTTCACC 3′).

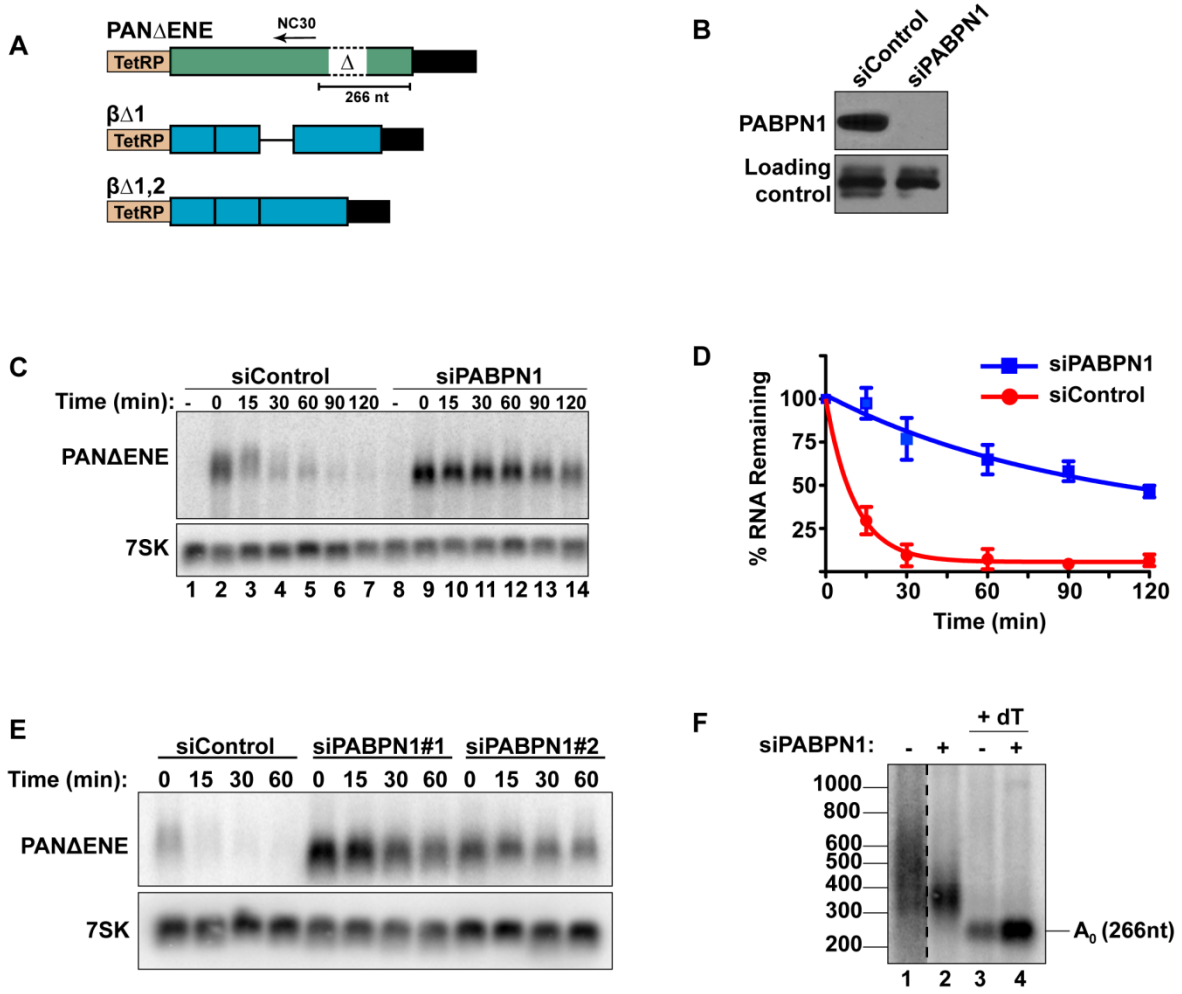

Figure 1. Rapid degradation of PANΔENE requires PABPN1. (**A**) Diagram of the TetRP-driven PANΔENE and β-globin reporters. PANΔENE has a deletion of the 79-nt ENE(Conrad et al., 2007). The position of oligonucleotide NC30 is shown (arrow). The β-globin reporters lacked either the first (βΔ1), or both (βΔ1,2), introns. (**B**) Western blot analysis of extracts from cells transfected with either a non-targeting control siRNA or a two-siRNA pool targeting PABPN1. Protein extract was probed with antibodies against either PABPN1 or PABPC1 (loading control). (**C**) Representative transcription pulse assay; the "-" lane samples were harvested prior to the two-hour transcription pulse. 7SK serves as a loading control (bottom). (**D**) Decay curves of the transcription pulse assay data (*n=3*). The data were fit to two-component exponential decay curves as described in the Materials and Methods. (**E**) Northern blot analysis of a transcription pulse-chase assay in cells transfected with a nontargeting control siRNA or either of two independent siRNAs against PABPN1. (**F**) Poly(A) tail length analysis of PANΔENE as described in text. Lane 1 is displayed at a darker exposure as indicated by dotted vertical lines.

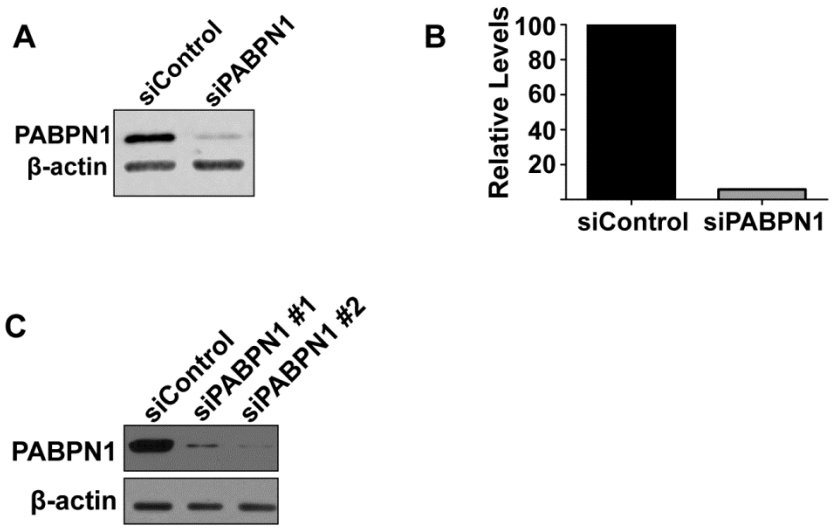

Figure 2. Related to Figure 1. (**A** and **B**) Quantitative western blot analysis of total lysate from cells transfected with either a control siRNA or an siRNA pool against PABPN1. The blot was probed with antibodies against PABPN1 and β-actin (loading control) and the signal was quantified relative to the β-actin loading control. (**C**) Western blot analysis of protein from the time zero samples in Figure 1E.

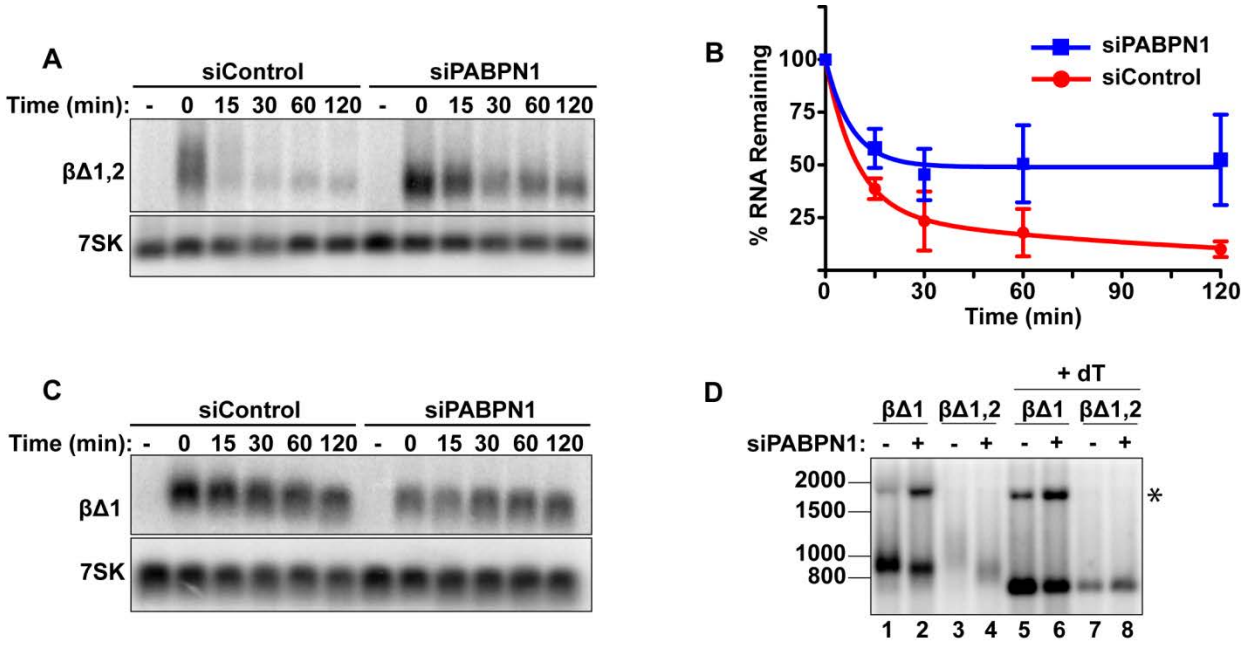

Figure 3. Rapid degradation of intronless β-globin requires PABPN1. (**A**) Representative transcription pulse assays using the intronless reporter (βΔ1,2) and cells transfected with the indicated siRNAs. (**B**) Decay curves of biological replicates of the transcription pulse assay data (*n=3*). (**C**) Representative transcription pulse assay of the spliced reporter (βΔ1). (**D**) Poly(A) tail analysis of βΔ1 and βΔ1,2. Time zero samples from cells transfected with the specified siRNAs was treated with RNase H and oligo(dT) as indicated. The β-globin pre-mRNA is denoted by an asterisk (*).

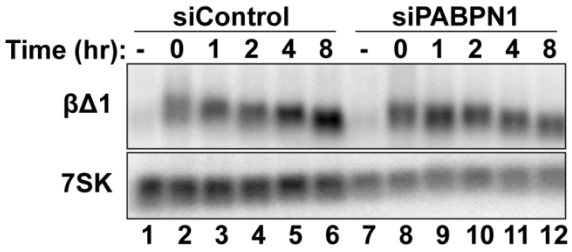

Figure 4. Related to Figure 3. Representative transcription pulse-chase of a spliced (βΔ1) reporter over an eight-hour time course.

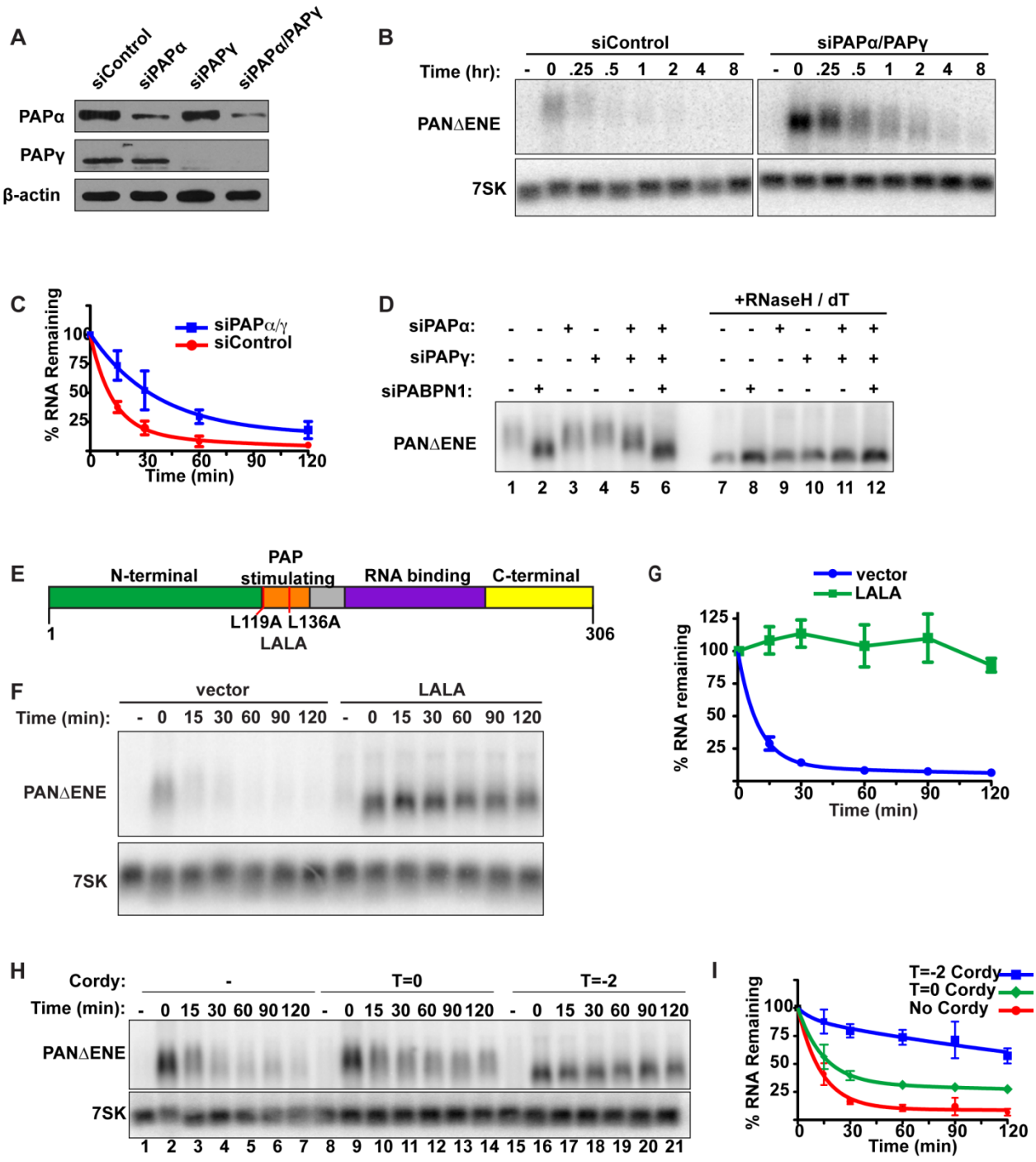

Figure 5. Polyadenylation is required for PANΔENE RNA decay. (**A**) Western blot analysis of cells transfected with either a control siRNA or siRNA pools targeting PAPα, PAPγ, or both. Specific proteins were detected using antibodies against PAPα, PAPγ, or β-actin (loading control). (**B** and **C**) Representative transcription pulse assay and decay curves of PANΔENE from cells transfected with the indicated siRNAs *(n≥4)*. (**D**) Length analysis of PANΔENE from cells transfected with the indicated siRNAs. RNA was harvested following a two-hour transcription pulse (T=0). The samples in lanes 7-12 were treated with RNase H and oligo(dT)

prior to northern blot analysis. (**E**) Schematic showing the domain structure of PABPN1 as well as the mutations made to generate the dominant negative construct, LALA. (**F**) Results from a transcription pulse from cells transfected with the dominant negative allele LALA. (**G**) Decay curves from a transcription pulse as shown in (**F)***(n≥3).* (**H** and **I**) Representative transcription pulse assay and decay curves of PANΔENE performed in the presence or absence of cordycepin *(n=3).* Cordycepin was added either during (T=-2) or after (T=0) the transcriptional pulse.

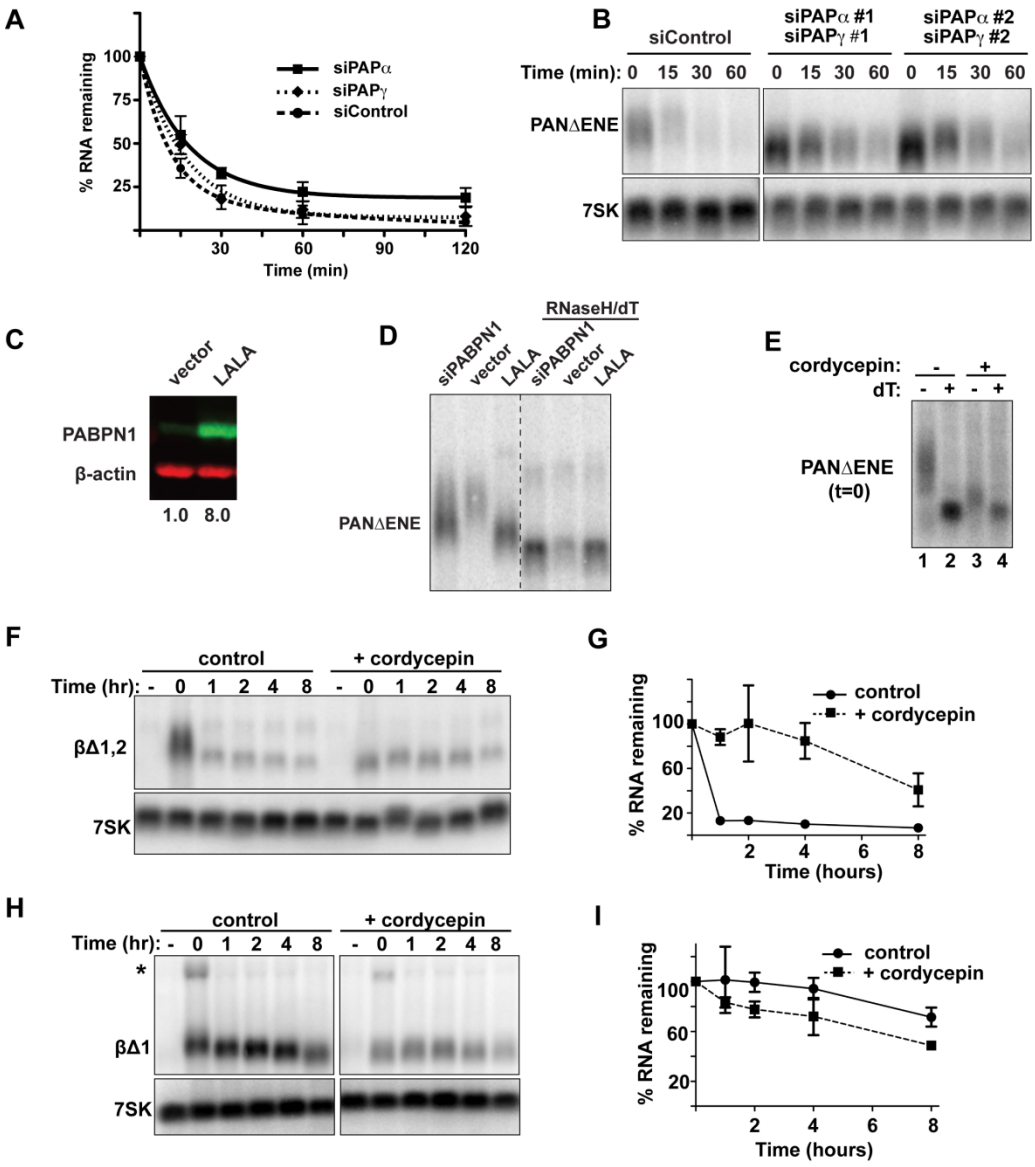

Figure 6. Related to Figure 5. (**A**) Decay curves of PANΔENE from cells transfected with the indicated siRNAs *(n=3)*. (**B**) Northern blot analysis of PANΔENE RNA from cells transfected with independent pools of siRNAs targeting PAPα and PAPγ (**C**) Quantitative western blot analysis of total lysate from cells transfected with either a vector control, or LALA PABPN1. The blot was probed with antibodies against either PABPN1 (green) or β-actin (red) as a loading control. The relative abundance of PABPN1 normalized to β-actin is indicated below each lane. Importantly, because these values include protein from untransfected cells, they are likely an underestimate of the degree of PABPN1 overexpression. (**D**) Relative poly(A) tail lengths of PANΔENE RNA following LALA overexpression. RNA from cells treated with siPABPN1 was included for comparison. (**E**) Relative poly(A) tail lengths of PANΔENE RNA following cordycepin treatment. PANΔENE RNA was cleaved with RNase H targeted by NC30 and oligo(dT) as indicated. The cleaved RNA was analyzed by northern blot using a 3´ end-specific probe. We estimated that the poly(A) tails in the presence of cordycepin were between ~20-50 nt.

(**F**) Representative transcription pulse assay of βΔ1,2 from cells induced in the presence or absence of cordycepin. (**G**) Linear interpolation of the results from (C)(*n=3*). (**H** and **I**) Same as in (**F**) and (**G**) except spliced β-globin (βΔ1) was assayed. The same blot and exposure are shown for both panels in (**H**).

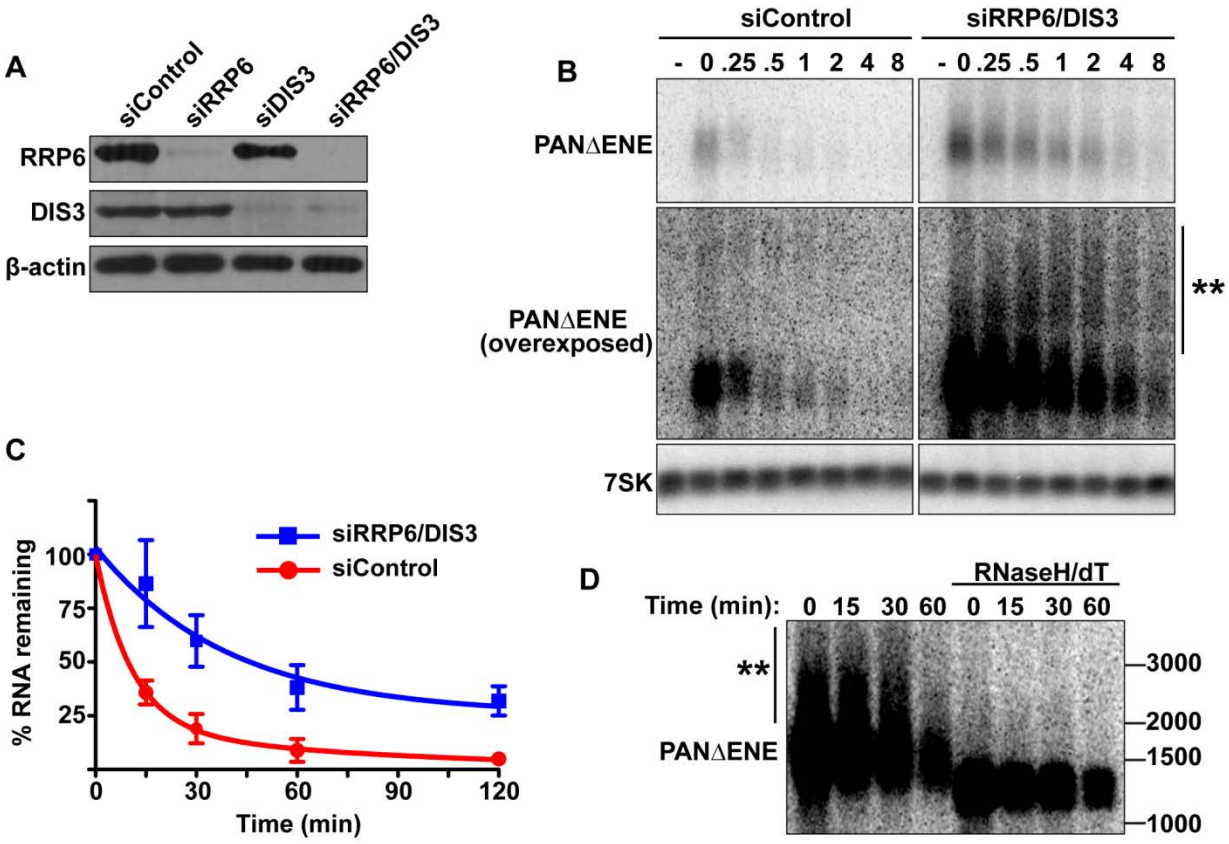

Figure 7. PANΔENE is stabilized and hyperadenylated upon exosome depletion. (**A**) Western blot analysis of cells transfected with either a non-targeting control siRNA or siRNA pools against RRP6, DIS3, or both. Blots were probed with antibodies specific for RRP6, DIS3, and β-actin (loading control). (**B**) *Top,* Representative transcription pulse assay from cells transfected with the indicated siRNAs. The same blot and exposure are shown for siControl and siRRP6/DIS3. *Lower panels,* Overexposed version of the blot and 7SK loading control. (**C**) Decay curves of the transcription pulse assays as shown in (**B**)(*n=3*). (**D**) Northern blot showing the mobility of PANΔENE RNA from RRP6 and DIS3 depleted cells before and after treatment with RNase H and oligo(dT).

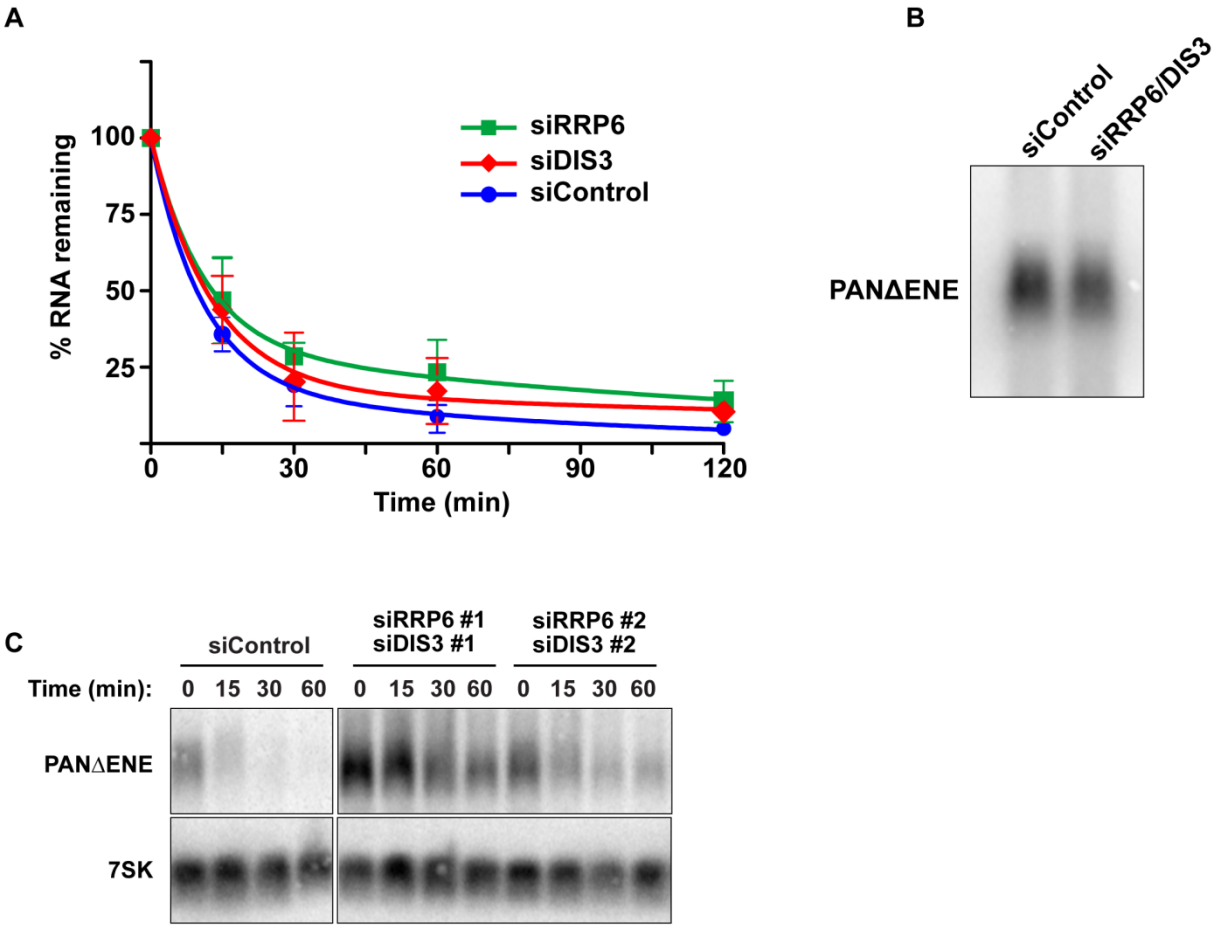

Figure 8. Related to Figure 7. **(A)** Decay curves of PANΔENE from cells transfected with the indicated siRNAs (*n=3)*. **(B)** A comparison of the relative lengths of PANΔENE from cells transfected with the indicated siRNAs. RNA is from the time zero samples (lanes 2 and 10 in Figure 7B). (**C**) Results from a transcript pulse assay of PANΔENE from cells transfected with two independent pools of siRNAs targeting RRP6 and DIS3.

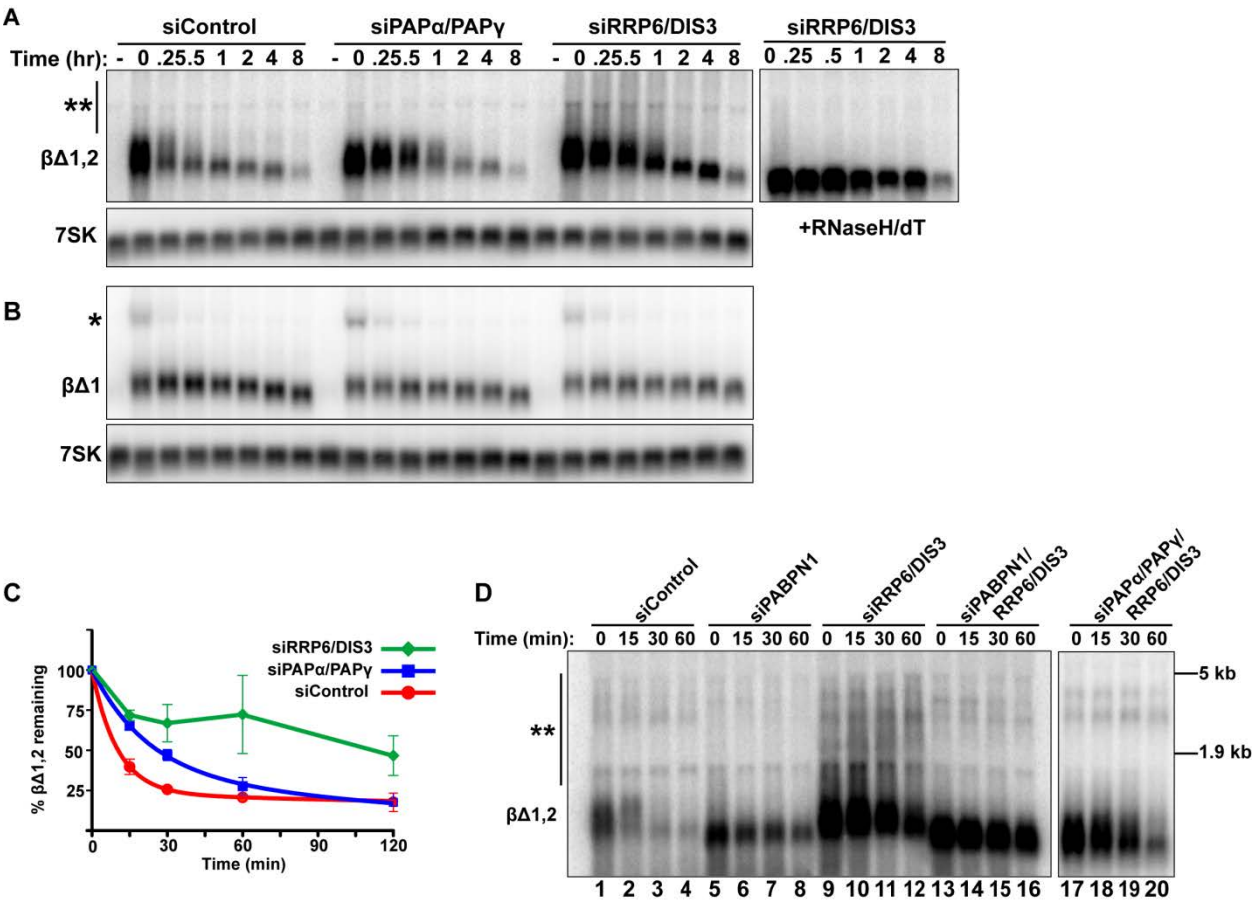

Figure 9. Intronless β-globin decay requires canonical PAP and exosome activity. (**A** and **B**) Representative blot of a transcriptional pulse assay with βΔ1,2 or βΔ1 reporters from cells transfected with the indicated siRNAs. *Right*, Samples from siRRP6/DIS3 lanes were treated with RNase H and oligo (dT). Single asterisk marks the β-globin pre-mRNA and double asterisk denotes the extremely hyperadenylated RNAs. (**C**) Decay curves of the intronless β-globin reporter data from cells transfected with the indicated siRNAs (*n=3*). The siDIS3/siRRP6 data were not well fit by two-component exponential decay regressions and are represented as linear interpolations (**D**) Results from a transcription pulse assay of βΔ1,2 from cells transfected with the indicated siRNAs. Blot was overexposed to reveal the extremely hyperadenylated RNAs (double asterisks). The relative mobility and sizes of the large and small ribosomal subunits are indicated on the right.

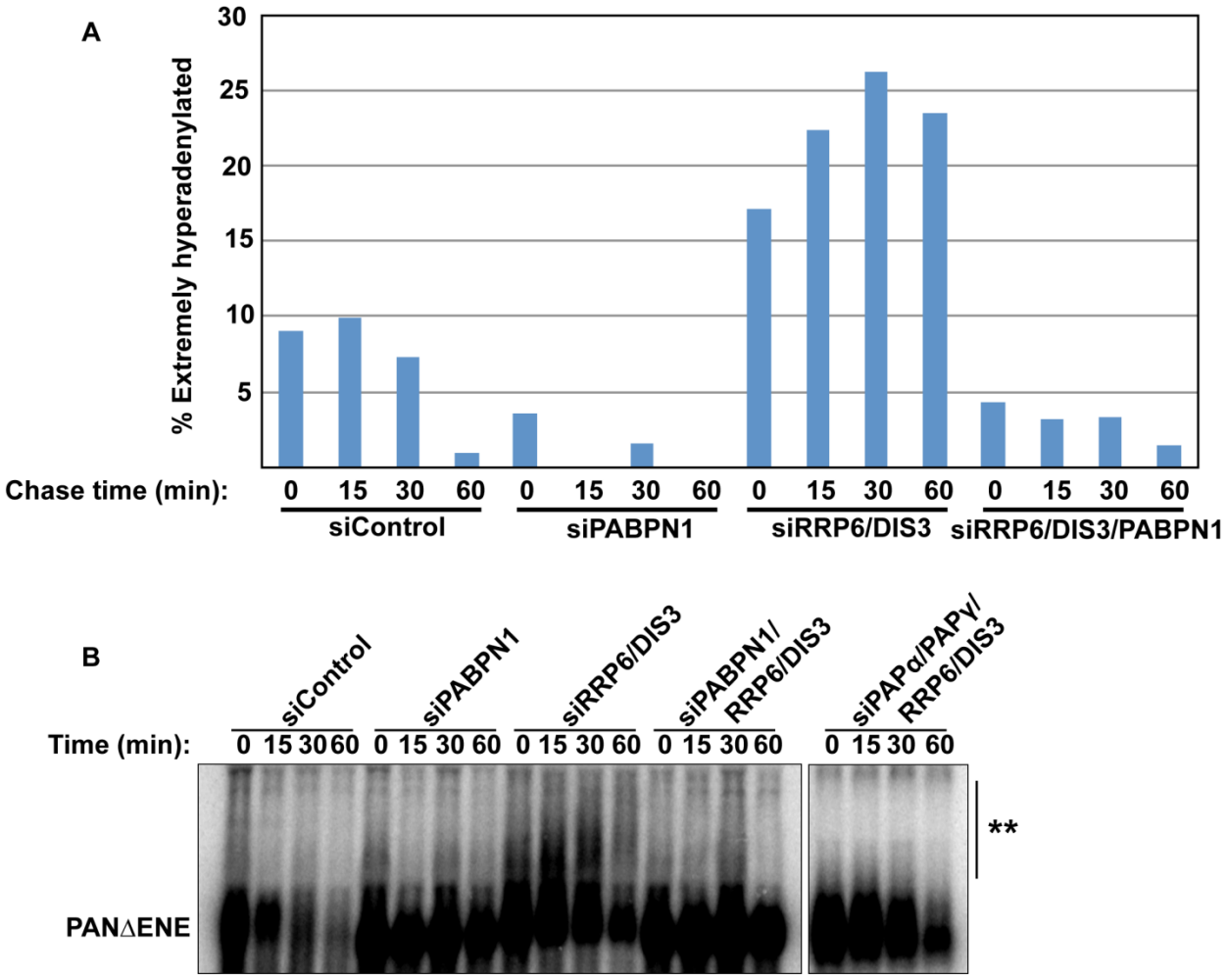

Figure 10. Related to Figure 9. (**A**) Quantitation of the percent of intronless β-globin which was "extremely hyperadenylated" in each lane of Figure 9D. The "% extremely hyperadenylated" was calculated by boxing the signal from each lane above the primary band, subtracting the background signal, and dividing by the total amount of signal in each lane. (**B**) Results from a transcription pulse assay of PANΔENE from cells transfected with the indicated siRNAs. Blot was overexposed to reveal the extremely hyperadenylated RNAs (double asterisks).

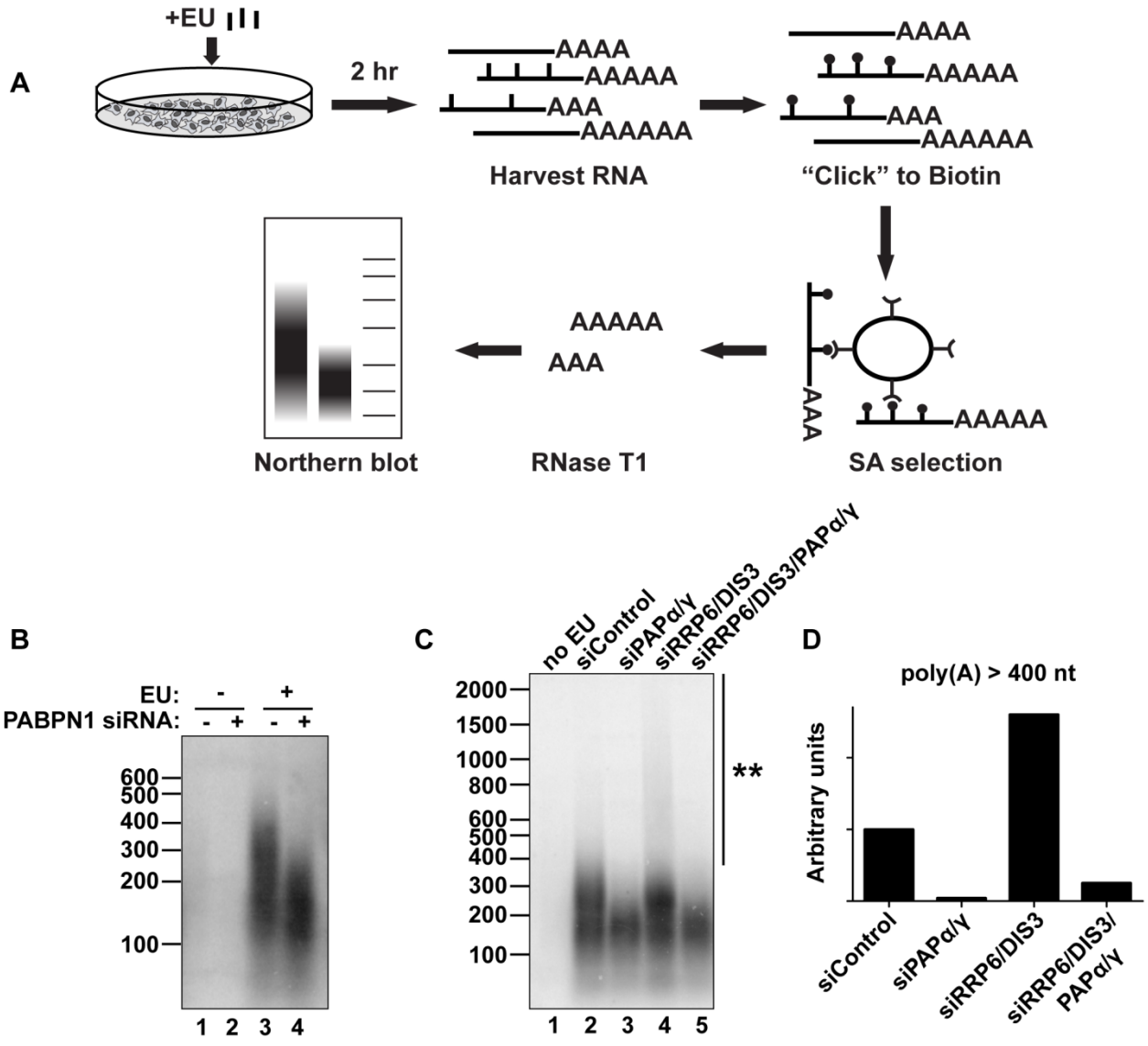

Figure 11. Endogenous transcripts are subject to PABPN1-mediated hyperadenylation. (**A**) Schematic diagram of the in vivo labeling procedure. See text for details. (**B**) Bulk poly(A) tail analysis of labeled RNAs from cells transfected with indicated siRNAs. Prior to harvesting the RNA, cells were incubated with EU for 2 hr as indicated. (**C**) In vivo labeling experiment with cells transfected with the indicated siRNAs. Double asterisk marks extensive hyperadenylation observed upon RRP6 and DIS3 co-depletion. (**D**) Quantification of the poly(A) tails greater than 400 nt in length from (**C**). We "boxed" signals corresponding to poly(A) tails >400 nt, subtracted background signal (no EU) and normalized to the signal from the siRNA control samples.

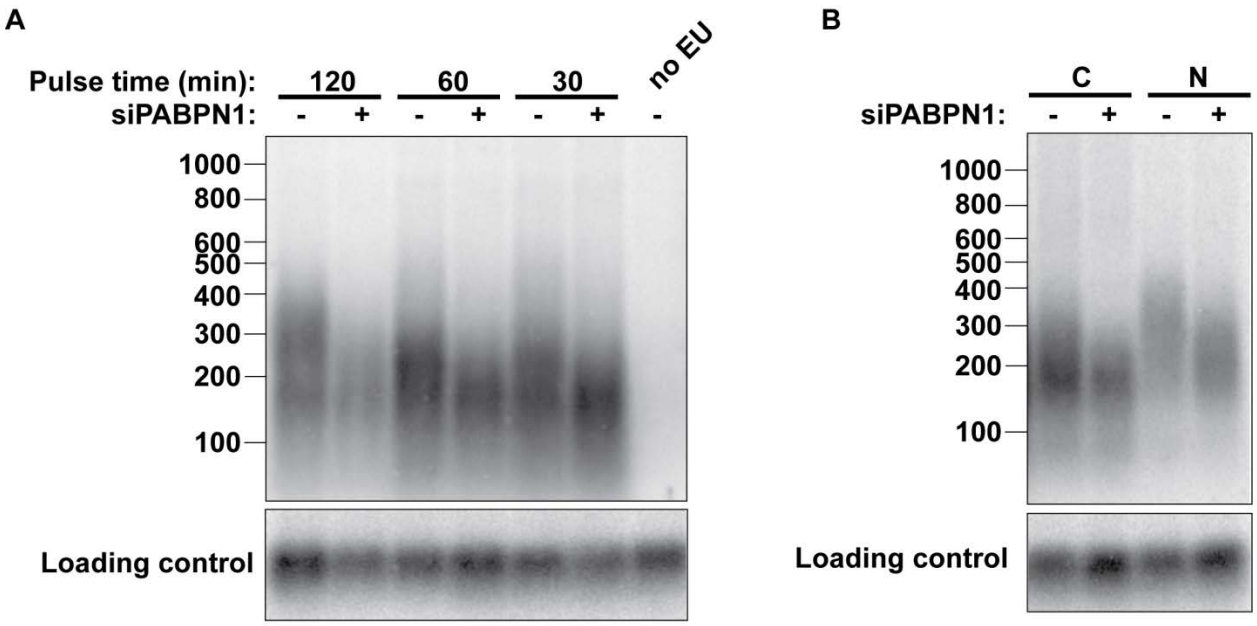

Figure 12. Related to Figure 11. (**A**) Bulk poly(A) tail analysis from cells transfected with either siControl or siPABPN1. Prior to harvesting the RNA, cells were incubated with EU for the indicated pulse times. To compensate for the shorter pulse times, increasing amounts of total RNA was used for the click reaction: 0.5 µg for the 120’ pulse, 1 µg for the 60’ pulse, and 2 µg for the 30’ pulse and the no EU sample. An exogenously added biotinylated DNA oligo (“Loading control”) was used to control for recovery. Note that this DNA only controls for RNA recovery and loading during the experimental procedure and does not necessarily reflect the total amounts of input RNA. (**B**) Nuclear/cytoplasmic distribution of poly(A) tails made in the presence or absence of PABPN1. Cells were separated into nuclear and cytoplasmic fractions following a two-hour EU pulse.

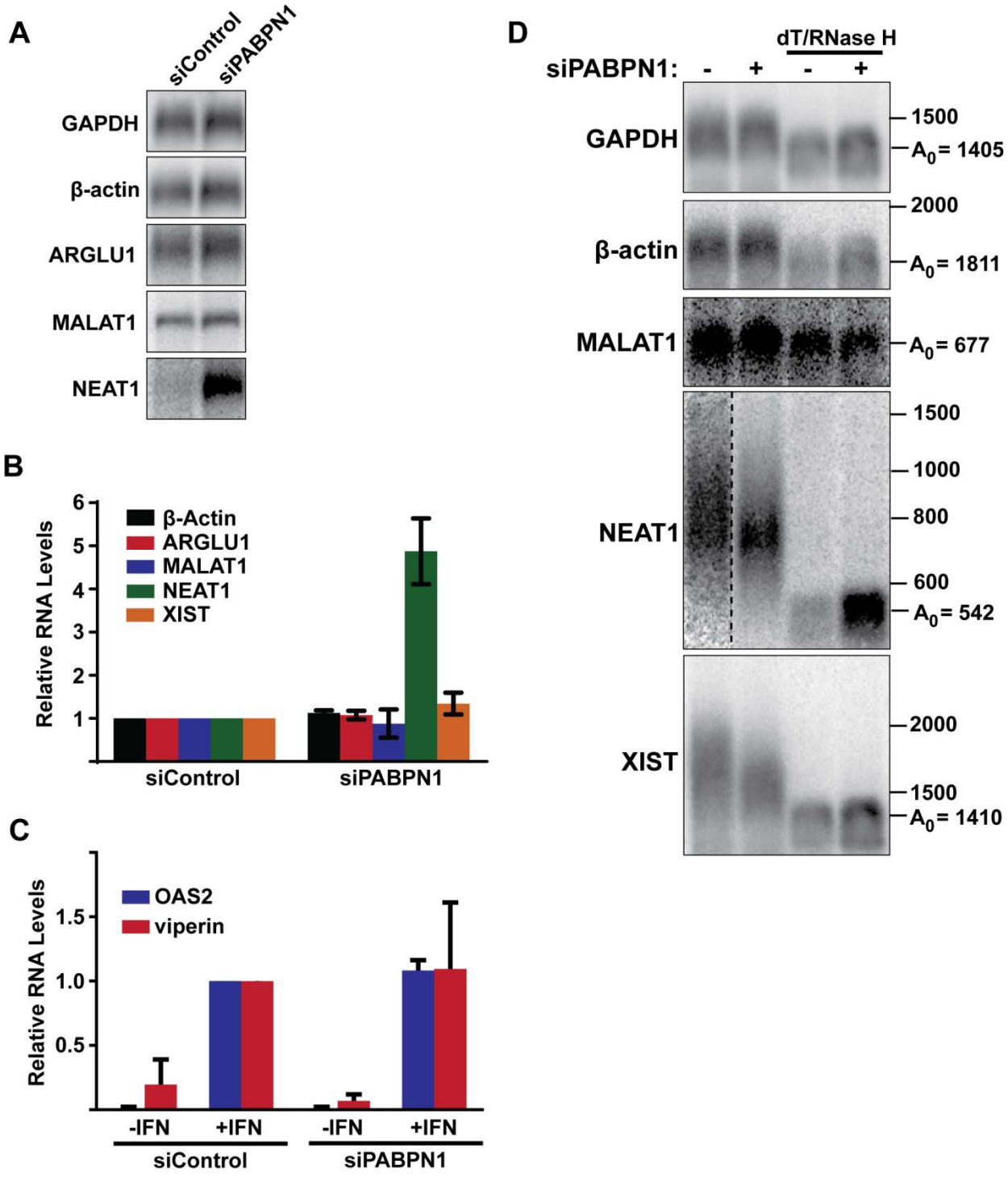

Figure 13. Effects of PABPN1 knockdown on specific endogenous RNAs. (**A**) Northern blot analysis of several endogenous lncRNAs and mRNAs after siRNA transfection. (**B**) Quantification of the relative abundance of endogeneous transcripts following PABPN1 depletion. All transcripts were quantified from northern blots shown in (A) except XIST, which was quantified from experiments shown in (D), because it is too large to detect by standard northern blot. GAPDH was used for normalization. (*n=3*) (**C**) Quantitative RT-PCR analysis

showing the relative induction of two interferon stimulated genes following PABPN1 depletion and the addition of interferon (IFN). Samples were quantified relative to the siControl/+IFN samples. GAPDH was used as a loading control (*n=3*). (**D**) Poly(A) tail length analysis of endogenous mRNAs and lncRNAs. Due to their large size, MALAT1, NEAT1, and XIST were cleaved with RNase H and a DNA oligo complementary to a region the near the 3´ end of each RNA. Samples were also treated with RNase H and oligo(dT) as indicated.

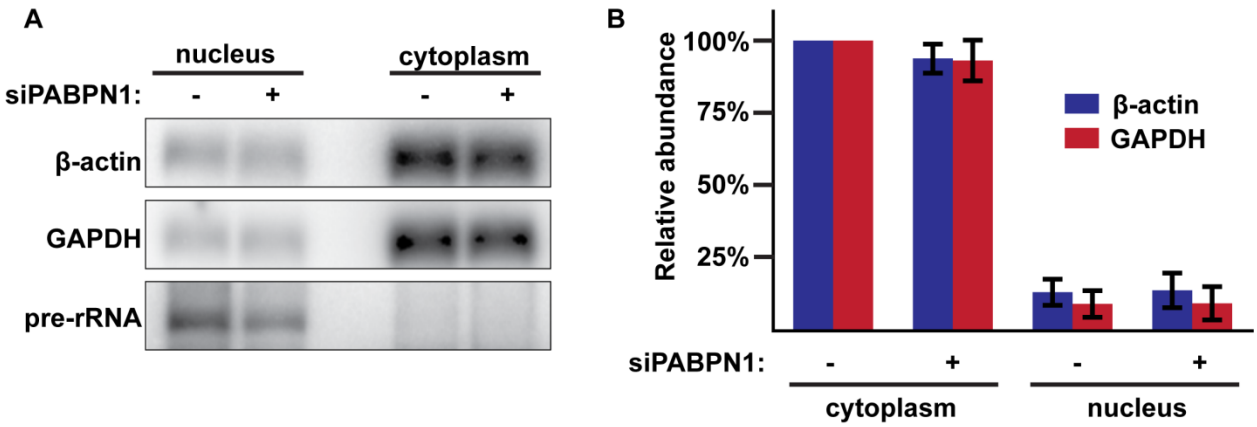

Figure 14. Related to Figure 13. (**A**) Northern blot analysis of nuclear and cytoplasmic fractions following PABPN1 depletion. The blot was probed for β-actin and GAPDH, as well as pre-ribosomal RNA in order to control for the quality of the fractionation. The amount of RNA loaded was kept constant at a 4:1 cytoplasmic:nuclear ratio. (**B**) Quantification of the results in panel (**A**) (*n=3*).

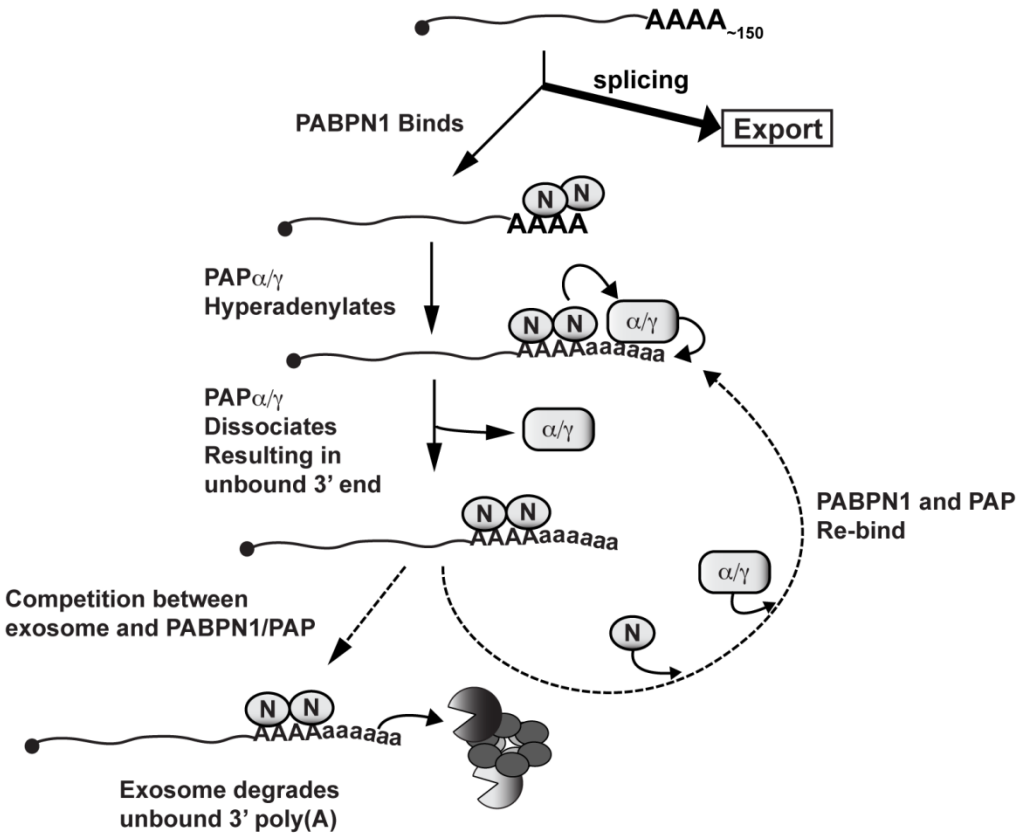

Figure 15. Model of PABPN1-mediated decay. (**A**) Diagram depicting a speculative mechanism of PABPN1-mediated decay. Symbols: PABPN1 (oval inscribed with "N"), PAPα/PAPγ (rectangle inscribed with "α/γ"), RRP6 and DIS3 (pacman symbols), core exosome (ring of ovals), lower case "a" in poly(A) tails denotes hyperadenylation. Note that PABPN1 is not depicted on the initial poly(A) tail for simplicity and is not meant to imply any mechanisms for PABPN1 in the initial polyadenylation event (see Discussion).

Table 1 Kinetic parameters estimated from regression analysis of decay data.

Reporter RNA	siRNA	% fast	$t_{1/2}$fast (min)	$t_{1/2}$slow (min)	R^2
PANΔENE	control*	94	7.4	>120	1.00
PANΔENE	siPABN1	6.6	15	116	0.96
βΔ1,2	control	73	6.4	89	1.00
βΔ1,2	siPABN1	51	5.4	>120	0.98
PANΔENE	control	84	8.4	65	1.00
PANΔENE	siPAPα/γ	43	15	63	0.99
PANΔENE	control	81	7.5	58	1.00
PANΔENE	siDIS3/RRP6	38	15	100	0.96
βΔ1,2	control	77	7	>120	1.00
βΔ1,2	siPAPα/γ	63	15	105	1.00
βΔ1,2	siDIS3/RRP6	ND	ND	ND	ND
RNA	**cordycepin**	**% fast**	**$t_{1/2}$fast (min)**	**$t_{1/2}$slow (min)**	**R^2**
PANΔENE	None	90	9.3	>2hr	1.00
PANΔENE	T=-2	10	5.9	>2hr	0.97
PANΔENE	T=0	67	10	>2hr	1.00
RNA	**Overexpression plasmid**	**% fast**	**$t_{1/2}$fast (min)**	**$t_{1/2}$slow (min)**	**R^2**
PANΔENE	Vector	88	6.6	>2hr	1.00
PANΔENE	YAFA	53	15	>2hr	0.93
PANΔENE	LALA	ND	ND	ND	ND

*siRNA controls were repeated for each new set of experiment. The control is listed in the row above the corresponding targeting siRNA.
ND: not determined

Table 2 siRNAs used in this study

siRNA	Sense sequence	Antisense sequence
PABPN1 #1	GGCCUUAGAUGAGUCCCUAtt	UAGGGACUCAUCUAAGGCCaa
PABPN1 #2	AGUCAACCGUGUUACCAUAtt	UAUGGUAACACGGUUGACUga
PAPα #1	GCCUCGACUUGUCUAUGGAtt	UCCAUAGACAAGUCGAGGCtg
PAPα #2	GUGCUGAUAUUGAUGCGUUtt	AACGCAUCAAUAUCAGCACct
PAPγ #1	CAGCUUAAAUGGUUGUAGAtt	UCUACAACCAUUUAAGCUGcg
PAPγ #2	CCAUAGAUGGGACUCCUAAtt	UUAGGAGUCCCAUCUAUGGat
RRP6 #1	GGAUCGAAGUAAAGUGACUtt	AGUCACUUUACUUCGAUCCtt
RRP6 #2	GAGUAUGAUUUUUACCGAAtt	UUCGGUAAAAAUCAUACUCat
DIS3 #1	GGAGCAUUACUGAAAAGGAtt	UCCUUUUCAGUAAUGCUCCag
DIS3 #2	CAAAAGCAAAGGAAUAGUAtt	UACUAUUCCUUUGCUUUUGaa

Chapter 3: PABPN1 promotes the degradation of snoRNA host genes

Introduction

3′ end formation is an essential and conserved feature of eukaryotic mRNA biogenesis. For most RNAs, 3′ end formation consists of two distinct steps: an endonucleolytic cleavage event followed by the addition of a poly(A) tail to the upstream cleavage product. In mammalian cells, both steps are dependent on the multiprotein cleavage and polyadenylation specificity factor (CPSF). Initially, CPSF cleaves the RNA 10-20 nt downstream of the AAUAAA cleavage and polyadenylation signal. Following cleavage, CPSF recruits the poly(A) polymerase (PAP) to the RNA through a direct physical interaction (Keller et al., 1991; Ruegsegger et al., 1996). This step is necessary because PAP has very weak affinity for RNA, and is essentially inactive on its own (Wahle et al., 1991). The physical interaction between PAP and CPSF tethers PAP to the transcript and allows polyadenylation to occur in a weakly processive manner. Binding of PABPN1 to the nascent oligo(A) tail further enhances polyadenylation activity (Wahle, 1991). Like CPSF, PABPN1 forms a physical interaction with PAP, and when both factors are present PAP is tightly tethered to the RNA and polyadenylation proceeds in a highly processive manner. Through a poorly understood mechanism, the interaction between CPSF and PAP is lost once the tail reaches a length of 200-300 nt, terminating processive polyadenylation (Kerwitz et al., 2003; Kuhn et al., 2009). For most RNAs, this is the end of 3' end processing. However, in some cases, polyadenylated transcripts are retained in the nucleus following 3′ end formation. These include the predominantly nuclear long non-coding RNAs (lncRNAs) as well as inefficiently exported mRNAs. Because these transcripts remain bound to PABPN1, they may continue to be slowly polyadenylated. For the purposes of this paper, this secondary polyadenylation is termed

“hyperadenylation”.

In eukaryotes, polyadenylation generally plays a positive role in gene expression. Nuclear-encoded poly(A) tails promote export, enhance translation, and inhibit RNA decay. Poly(A) tails serve a very different function in bacteria, where polyadenylation is an important first step in RNA degradation (reviewed in (Mohanty and Kushner, 2011)). Poly(A) tails in bacteria are thought to function as an unstructured "landing pad", allowing the decay machinery easier access to the 3' end of the RNA. Recently, polyadenylation has also been shown to facilitate RNA turnover in eukaryotes. The first such pathway to be discovered involves the Trf4-Air2-Mtr4 polyadenylation (TRAMP) complex, which is required for the degradation of a variety of nuclear RNAs, including aberrant rRNA, tRNA, snoRNAs, snRNAs, lncRNAs, and mRNAs (LaCava et al., 2005; Schmidt and Butler, 2013; Vanacova et al., 2005; Wyers et al., 2005). TRAMP's polyadenylation activity is supplied by the non-canonical poly(A) polymerase Trf4, which catalyzes the addition of a short oligo(A) tail to the 3' end of target transcripts. These RNAs are then targeted for decay by the nuclear exosome, a highly conserved protein complex consisting of a nine-subunit core that associates with the Rrp6 and Dis3/Rrp44 exonucleases (Schneider and Tollervey, 2013). Despite the importance of TRAMP for exosome-mediated decay, the role that polyadenylation plays in this process remains unclear. Consistently, multiple studies have found that polyadenylation is not strictly necessary for exosome activity, either in vitro or in vivo (Callahan and Butler, 2010; Houseley et al., 2007; Rougemaille et al., 2007; San Paolo et al., 2009). In one particularly striking example, a catalytic site mutant in Trf4 is able to rescue the steady-state levels of ~90% of decay targets in a trf4Δ strain (San Paolo et al., 2009). These observations suggest that polyadenylation may only be required for a subset of TRAMP targets.

Recently, the fission yeast nuclear poly(A)-binding protein Pab2 was shown to stimulate exosome activity toward a variety of target RNAs (Chen et al., 2011; Lemay et al., 2010; Lemieux et al., 2011; Yamanaka et al., 2010). Like TRAMP, Pab2-mediated decay is associated with polyadenylation. In addition to the poly(A) tail, decay requires the CPSF-component Rna15 and the poly(A) polymerase Pla1 (Yamanaka et al., 2010). Even so, the role of polyadenylation is still unclear. Pab2 does not stimulate the activity of its cognate PAP in vitro (Eckmann et al., 2011) and Pab2 deletion results in longer poly(A) tails at the bulk level and for at least some targets (Lemay et al., 2010; Perreault et al., 2007). Importantly, Pab2 physically interacts with Dis3 and Rrp6 (Lemay et al., 2010). These observations suggest Pab2 promotes decay by direct recruitment of the exosome, rather than through polyadenylation. In this model, the poly(A) tail is necessary to provide a binding site for Pab2, but does not directly stimulate decay.

Pab2 shares 47% identity and 66% similarity with the human protein PABPN1 (Perreault et al., 2007). Recently, we and others showed that PABPN1 stimulates the degradation of lncRNAs, a viral noncoding RNA, and a nuclear-retained mRNA in human cells (Beaulieu et al., 2012; Bresson and Conrad, 2013). As discussed above, PABPN1 has a well-defined role in polyadenylation. Our work with the KSHV polyadenylated nuclear (PAN) RNA suggests that PABPN1-mediated polyadenylation is essential for decay. PAN RNA is a noncoding capped and polyadenylated RNA which accumulates to high levels in the nucleus. Nuclear accumulation depends on the presence of a 79 nt stability element termed the ENE, which interacts in cis with the poly(A) tail of the transcript. Deletion of the ENE results in a highly unstable transcript ($t_{1/2}$ < 15 minutes) which is degraded by a pathway involving PABPN1, the canonical poly(A) polymerases PAPα and PAPγ, and the nuclear exosome (Bresson and Conrad, 2013; Conrad et al., 2006). We have termed this pathway PABPN1-PAP-mediated decay (PPD). Several lines of

evidence suggest that PANΔENE turnover is polyadenylation dependent (Bresson and Conrad, 2013). First, transient poly(A) tail extension is observed prior to decay, and this phenotype is exacerbated in exosome-depleted cells. Second, the polyadenylation inhibitor cordycepin prevents poly(A)-tail extension and stabilizes PANΔENE. Finally, PAP-stimulating mutations in PABPN1 completely abolish decay. While these observations strongly support a role for polyadenylation in PANΔENE RNA decay, it remains unclear if polyadenylation is a general feature of PABPN1-mediated decay.

In this paper, we further investigate the role of polyadenylation in PABPN1-mediated decay. We provide evidence that decay is stimulated by the hyperadenylation phase of canonical polyadenylation, and that poly(A) tail extension is necessary but not sufficient to promote decay. In order to identify the subset of PABPN1-targets that are degraded in a polyadenylation-dependent manner, we also perform RNAseq analysis on cells expressing a polyadenylation-defective PABPN1 mutant or cells depleted of PAP. Our results show that polyadenylation appears to be generally required for PABPN1-mediated decay.

Results

CPSF-mediated polyadenylation activity is not required for decay

In order to measure PANΔENE RNA stability, we employed a well-characterized transcription pulse-chase strategy (Bresson and Conrad, 2013; Sahin et al., 2010). In this assay, 293A cells stably expressing a tetracycline-responsive transcriptional activator (293ATOA) are transfected with a plasmid containing PANΔENE under the control of a tetracycline-responsive promoter. Transcription is repressed in the presence of the tetracycline analog doxycycline (dox), and activated following washout. In our experiments, we performed a two hour transcriptional

pulse, and then readded dox to the media to repress transcription. Time points were collected following transcription shutoff, and PANΔENE stability was monitored by northern blot.

In vitro, both PABPN1 and CPSF physically interact with PAP and stimulate polyadenylation activity (Kerwitz et al., 2003). CPSF participates in the polyadenylation reaction until the tail reaches a length of 200-300 nt and beyond that length, polyadenylation is dependent solely on PABPN1. PABPN1-mediated polyadenylation is required for the degradation of PANΔENE, but we have not previously assessed a requirement for CPSF's polyadenylation activity. To address this question, we devised a way to generate cleaved and polyadenylated PAN RNA while bypassing the requirement for CPSF (Figure 16A). Our approach took advantage of the unusual processing of the MALAT1 lncRNA. The MALAT1 3′ end is generated by RNase P, which cleaves directly upstream of a tRNA-like element within the body of the transcript (Wilusz et al., 2008). We cloned this element into ENE-lacking PAN and placed a short 35 nt poly(A) tail directly upstream (PANΔENE-A_{35}). Processing was highly efficient, as we were unable to detect any uncleaved transcript (data not shown). Once the RNA is cleaved by RNase P, the genomically encoded poly(A) tail is exposed and was extended. Importantly, because the cleaved transcript lacks an AAUAAA site, any further polyadenylation is dependent on PABPN1 and PAP alone. This mimics the hyperadenylation phase of canonical polyadenylation, in which polyadenylation is dependent on PABPN1 but not CPSF. We measured the steady state poly(A) tail to be ~100-500 nt in length (Figure 16B). Next, we compared the stability of RNase P-processed PANΔENE to CPSF-processed PANΔENE (PANΔENE-AAUAAA). Notably, both transcripts had equivalent decay kinetics (Figure 16C, D), suggesting that CPSF-mediated polyadenylation is not required for decay.

To confirm that PANΔENE-A_{35} is in fact degraded by PPD, we measured its stability following PABPN1 depletion (Figure 16E, F). As expected, PANΔENE-A_{35} was stabilized in the absence of PABPN1. Taken together, these results suggest that PABPN1-mediated polyadenylation is necessary to stimulate decay while CPSF-mediated polyadenylation is dispensable.

Polyadenylation is not sufficient for decay

PABPN1 stimulates polyadenylation by physically recruiting PAP to the poly(A) tail (Kerwitz et al., 2003). PAP recruitment is necessary to trigger decay of PANΔENE (Bresson and Conrad, 2013), but it is not known if this step is sufficient. To address this question, we employed a tethering assay to recruit PAP to PAN RNA in the absence of PABPN1.We inserted six MS2 binding sites into PANΔENE to promote tethering of MS2-fusion proteins to a position just upstream of the poly(A) tail (PAN-6MS2) (Figure 17A). When coexpressed with MS2 alone, PAN-6MS2 was rapidly degraded in siControl treated cells (Figure 17B, lanes 5-8), but stable in siPABPN1 treated cells (Figure 17B, lanes 13-16). This shows that the MS2 hairpins do not interfere with decay, and that PAN-6MS2 is normally degraded by PPD. Next, we coexpressed PAN-6MS2 with MS2-PAP. As expected, PAN-6MS2 was rapidly degraded in control cells (Figure 17B, lanes 1-4). However, when coexpressed in a PABPN1-depleted background, MS2-PAP was not able to rescue decay, despite promoting robust polyadenylation (Figure 17B, lanes 9-12). We conclude that polyadenylation alone is not sufficient to stimulate PAN RNA decay.

Global analysis of PPD targets

While active polyadenylation is necessary for the degradation of some PABPN1 targets (i.e. PANΔENE and intronless β-globin), it is not clear if this is generally true. Thus, we next aimed to identify the set of PABPN1-targets that are degraded in a polyadenylation-dependent manner. To this end, we created a stable cell line expressing PABPN1-LALA (hereafter, "LALA") under control of the tetracycline promoter (see Materials and Methods). LALA contains two point mutants (L119A and L136A) that render the protein unable to stimulate polyadenylation by PAP. LALA retains RNA binding activity and acts as a potent dominant negative inhibitor of both polyadenylation (Kuhn et al., 2009) and PANΔENE RNA decay (Bresson and Conrad, 2013). Following a three day induction of LALA expression, we collected RNA in preparation for high-throughput sequencing. Under these conditions, LALA was expressed at a roughly 1:1 ratio with endogenous, wild type PABPN1 (data not shown). We also prepared RNA from cells depleted of PAPα and PAPγ (collectively, "PAP"), the polymerases required for PABPN1-mediated decay of PANΔENE. By comparing the overlap in upregulated genes between LALA and siPAP knockdown, we aimed to generate a high-confidence list of polyadenylation-dependent decay targets. For each condition (control, LALA, and siPAP), we prepared both total RNA and a nuclear-enriched fraction. Each sample was submitted in duplicate for Illumina-based RNAseq.

Sequencing reads were aligned to the human genome using Bowtie. Gene expression was calculated based on the number of reads mapping to each gene, and normalized using RPKM, which accounts for both the length of the gene and the total number of sequencing reads in the sample. We defined differentially expressed genes (DEGs) as significantly changed genes ($p<.05$) with a false discovery rate of less than 5%. A global analysis of one set of DEGs, comparing nuclear enriched fractions from LALA and control, is shown in Figure 18A. Not

surprisingly, most DEGs are upregulated, consistent with PPD's role in RNA decay. Next, we compared the overlap in significantly upregulated genes between each dataset (Figure 18B). As expected, comparisons between the total and nuclear-enriched fractions for each condition showed strong overlap. More interestingly, we also observed substantial overlap between the upregulated DEGs when comparing the LALA and siPAP datasets. We compiled a list of genes, 95 in total, that were upregulated in all four data sets to generate a high confidence list of potential polyadenylation-dependent decay targets. Of these 95 genes, a slight majority were mRNAs, while the rest were various noncoding RNAs, including lncRNAs (33%), small nucleolar RNA (snoRNA) host genes (5.3%), and pseudogenes (5.3%) (Figure 18C).

snoRNA host genes are targets of PPD

We decided to further analyze the snoRNA host genes that were upregulated in our RNAseq analysis. snoRNAs guide the posttranscriptional modification of rRNAs and other RNAs, while small cajal body RNAs (scaRNAs) control the modification of spliceosomal RNAs. In mammalian cells, most snoRNAs and scaRNAs are encoded within the introns of protein coding genes, but a small number are found within the introns of noncoding host genes (Rearick et al., 2011). There are currently 24 annotated noncoding snoRNA host genes (SNHG1 through SNHG24), which collectively contain 83 snoRNAs and scaRNAs (Table S1). Nineteen SNHGs were expressed in our cell line, and 5 (26%) were significantly upregulated in all four datasets. This is consistent with a prior study in which three SNHGs were identified in an RNAseq analysis following PABPN1 depletion (Beaulieu et al., 2012). While annotating our dataset, we also identified a potentially novel snoRNA host gene. RPL32P3 was significantly upregulated in

three of four datasets, and contains a predicted snoRNA (SNORA7B) in the first intron. Because RPL32P3 lacks any significant open reading frame, it meets the definition of a noncoding snoRNA host gene. We have tentatively renamed it SNHG25 and included it in our analysis.

A number of other SNHGs were also upregulated, but did not meet our stringent cutoffs for inclusion. In order to obtain a more complete picture of which SNHGs were targets of PPD, we performed RT-qPCR on all 20 expressed SNHGs following PABPN1 depletion, PAP depletion, or LALA expression. The results are represented by the heatmap in Figure 19A (see figure legend for details). Eleven of the 20 genes (55%) were significantly upregulated following PABPN1 depletion. Notably, these same 11 genes were also increased following PAP depletion or LALA expression (Figure 19A), suggesting that polyadenylation is essential for PABPN1-mediated decay of snoRNA host genes.

The increase in spliced SNHG levels could be explained by enhanced stability prior to splicing. To explore this possibility, we compared the relative changes in spliced and unspliced transcripts following inhibition of PPD (Figure 19B). In most cases, the unspliced transcript was either unchanged or only modestly upregulated, implying that PPD primarily targets the spliced transcript. The major exception was SNHG25, for which both the unspliced and spliced versions were significantly upregulated.

In principle, the increase in SNHG levels could reflect changes in transcription rather than decay. To exclude this possibility, we performed a nuclear run-on assay to directly assess transcription rates for several snoRNA host genes. In a nuclear run-on assay, purified nuclei are harvested and chilled, effectively “freezing” elongating polymerases in place. Transcription is resumed upon addition of fresh nucleotide triphosphates (NTPs) and incubation at 30°C.

Importantly, the reaction conditions prevent new transcription initiation, so only previously engaged polymerases will incorporate NTPs. We performed our run-ons in the presence of 4-thioUTP (Davidson et al., 2012), a modified nucleotide which is readily incorporated by RNA polymerase into nascent transcripts. Subsequently, 4-thioUTP labeled RNA can be purified with thiol-specific biotinylation and selection on streptavidin-coated magnetic beads. All of our experiments included a "no 4-thioUTP" control to account for nonspecific pulldown of unlabeled RNAs. As shown in Figure 19C, only a small percentage of the total signal was due to nonspecific pulldown (compare column 1 with column 2 and 3 in each bar graph). Notably, transcription rates for SNHG9 and SNHG19 were decreased 20-40% following PABPN1 depletion, despite an 8-fold and 6-fold increase in steady state levels, respectively (Figure 19A). This rules out a transcriptional role for PABPN1/PAP in SNHG levels.

To confirm that PABPN1/PAP act posttranscriptionally, we measured RNA stability directly. We performed a transcription pulse-chase assay using 4-thiouridine nucleoside (4SU), which is efficiently incorporated into newly made transcripts (Zeiner et al., 2008). As with the nuclear run-ons, the addition of 4SU allows us to attach a biotin tag to labeled transcripts, and then select using streptavidin-coated beads. In this specific experiment, we treated cells with either a control siRNA or siRNAs targeting PABPN1. Following knockdown, cells were pulsed for one hour in the presence of 4SU, after which the cells were washed and grown in label-free media for an additional hour (Figure 20A). This step was necessary to allow unincorporated 4SU within the cell to be used up. After the one hour washout step, we collected the T=0 time point and subsequent time points 30', 60', and 120' later. For SNHG9 and SNHG19, PABPN1 depletion significantly impaired decay (Figure 20B, C). In contrast, there was no effect on the degradation of introns (Figure 20D, E) or c-myc (Figure 20G), an unstable cytoplasmic mRNA.

Similarly, there was no effect on GAPDH mRNA, which was completely stable throughout the time course (Figure 20F).

Nuclear retention influences susceptibility to PPD

Most SNHGs contain numerous stop codons upstream of the final exon-exon splice junction, and as such are potential candidates for nonsense mediated decay (NMD) (Chang et al., 2007). Indeed, both SNHG1 and SNHG2/GAS5 are known NMD substrates (Smith and Steitz, 1998; Tycowski et al., 1996). To test whether other SNHGs might also be targeted by NMD, we treated cells with the translation inhibitor cycloheximide for six hours. The results are summarized in Figure 19A. Of the 20 expressed SNHGs, we identified 9 (45%) as potential NMD targets. Surprisingly, most transcripts seem to be targeted by either PPD or NMD, but not both. The exceptions to this are SNHG1 and SNHG20, and possibly SNHG9. Three transcripts (SNHG5, SNHG6, and SNHG16) are not degraded by either pathway.

We next examined the effects of inactivating both decay pathways simultaneously. We hypothesized that NMD may serve as a backup pathway to PPD. SNHGs which evade PPD-mediated decay in the nucleus would then be degraded by NMD in the cytoplasm. To test this hypothesis, cells were treated with either a control siRNA or siRNAs targeting PAP, and following knockdown, either with or without cycloheximide. SNHG levels were assessed by RT-qPCR. To our surprise, PAP knockdown and cycloheximide treatment did not show any additive effects beyond PAP knockdown alone (Figure 21A). These results suggest that NMD does not simply degrade SNHGs which escape PPD; rather, each decay pathway targets a distinct set of SNHGs.

Why are some SNHGs degraded by NMD and others by PPD? Considering NMD targets cytoplasmic transcripts, and PPD targets nuclear transcripts, we hypothesized that differences in RNA localization could be responsible. To test this hypothesis, we derived a "nuclear enrichment" score for each of the 13,114 genes with RPKM > 0.5 in our dataset. Using the control datasets, we calculated a given gene's nuclear enrichment by dividing $RPKM_{nuclear}$ by $RPKM_{total}$. Genes with higher values are more enriched in the nucleus. Importantly, these values only allow us to rank genes by their relative nuclear enrichment; we cannot directly compute what percentage of a gene's RNA is nuclear versus cytoplasmic. In Figure 2B, we plotted the nuclear enrichment scores for all 13,114 genes (see figure legend for details). Next, we plotted the RT-qPCR data from Figure 19A by $\log_2$(fold change), and colored each SNHG by its nuclear enrichment value (Figure 21C). Strikingly, PPD was highly correlated with nuclear enrichment. PPD targets were typically strongly nuclear, while non-PPD targets tended to be more cytoplasmic. We conclude that differences in nuclear retention influence the susceptibility of each SNHG to PABPN1/PAP decay.

Discussion

In this paper, we further explored the role of polyadenylation in PABPN1-mediated decay. We show that PANΔENE decay is stimulated by hyperadenylation, but this is not sufficient for degradation. In order to identify polyadenylation-dependent decay targets, we performed RNAseq using human cells defective in polyadenylation, either through expression of a dominant-negative PABPN1 protein, or through depletion of poly(A) polymerase. We identified a variety of RNA classes, including snoRNA host genes, which we examined in detail.

The role of polyadenylation in PABPN1-mediated decay

Polyadenylation in mammalian cells consists of two phases. In the initial phase, polyadenylation depends on both PABPN1 and CPSF. Together, these two factors tightly tether PAP to the 3′ end of the RNA, resulting in highly processive polyadenylation. This phase continues until the tail reaches a length of 200-300 nt, at which point the interaction between CPSF and PAP is lost. In the second phase, polyadenylation is dependent solely on PABPN1. Because PAP is not as tightly tethered to RNA in the absence of CPSF, polyadenylation is only weakly processive in this phase. Here, we showed that CPSF is not required for PPD, as an artificially polyadenylated transcript lacking a CPSF site is degraded at the same rate as wild type PAN (Figure 16). These observations suggest decay is primarily triggered by the second "hyperadenylation" phase of canonical polyadenylation.

To test whether hyperadenylation is sufficient to stimulate decay, we artificially tethered PAP to PANΔENE RNA. PAP tethering failed to rescue decay in the absence of PABPN1 (Figure 17), arguing that polyadenylation is not sufficient to drive decay. These results suggest that PABPN1 stimulates decay via at least two mechanisms: 1) stimulation of polyadenylation, and 2) a polyadenylation-independent mechanism, perhaps direct recruitment of the exosome complex. Notably, the *S. pombe* homolog of PABPN1 appears to stimulate decay through physical recruitment of Rrp6 and Dis3 to target RNAs (Lemay et al., 2010; Yamanaka et al., 2010). Our results suggest that this ancestral mechanism may also be conserved in humans.

snoRNA host gene decay

snoRNAs are involved in the modification of ribosomal RNA, while their structural homologs scaRNAs are required for the modification of spliceosomal RNAs. In mammalian cells the vast majority of snoRNAs and scaRNAs are encoded within the introns of other genes. Following splicing and lariat debranching, snoRNA-containing introns are exonucleolytically processed from both ends, releasing the mature snoRNA (Kiss and Filipowicz, 1995). Most intronic snoRNAs reside in protein coding genes, many of which have roles in translation and ribosome biogenesis. It has been argued that cotranscription of snoRNAs with ribosomal protein genes allows the cell to coregulate the expression of the many different factors required for ribosome biogenesis. A significant fraction of snoRNAs (~25% in humans) are also found in the introns of noncoding host genes. These RNAs show very little conservation within exonic regions, suggesting they have no function apart from snoRNA biogenesis (but see (Askarian-Amiri et al., 2011; Kino et al., 2010) for two exceptions). Consistent with this idea, spliced SNHGs also tend to be highly unstable.

Noncoding SNHGs were previously thought to be degraded primarily by nonsense mediated decay. Recently however, three SNHGs were shown to be upregulated following PABPN1-depletion (Beaulieu et al., 2012), suggesting that at least some SNHGs might be degraded by alternative pathway. Here we demonstrated that the majority of SNHGs are degraded in a PABPN1-PAP dependent manner (Figure 22). Moreover, our experiment involving the polyadenylation-defective mutant of PABPN1 (LALA), show that degradation requires active poly(A) tail extension and not simply a poly(A) tail. We also identified several potential NMD targets, in addition to the two that were previously reported. Importantly, we showed that NMD and PPD are largely non-overlapping with respect to SNHG decay, with each degradation pathway targeting a distinct subset of SNHGs (Figure 19). Target specificity is at

least partially influenced by RNA localization: more nuclear SNHGs are degraded by PPD, and more cytoplasmic SNHGs are degraded by NMD (Figure 21). Further study is needed to elucidate the mechanism by which some SNHGs are retained in the nucleus following splicing.

Materials and Methods

Plasmids

The TetRP-driven PANΔENE-AAUAAA was generated by digestion of PANΔ115-Bgl (Conrad et al., 2006) with NcoI and EcoRV. The resulting fragment was cloned into TRP-PANΔ79 (Conrad et al., 2007) digested with the same enzymes. TetRP-driven PAN PANΔENE-AAUAAA includes a 115 nt deletion encompassing the 79 nt ENE. In order to generate PANΔENE-A35, we performed PCR on genomic DNA to amplify a stretch of the Malat1 gene. We cloned the mascRNA element along with 60 nucleotides of downstream sequence in case these were essential for RNase P cleavage. The forward primer included a stretch of 35 adenosine residues in order to generate a short poly(A) tail directly upstream of the predicted cleavage site in the mature RNA. The resulting PCR fragment was cloned into PANΔ115-AAUAAA using the BglII site.

PABPN1-LALA was previously described (Bresson and Conrad, 2013). To generate the stable cell line expressing LALA, we used the pcDNA5-FRT/Flp recombinase system (Invitrogen), which allows integration of a given sequence into the genome. We first replaced the CMV promoter of pcDNA5-FRT with TetRP (pcDNA5-FRT-TRP). Next, we digested PABPN1-LALA with BamHI and NotI and cloned the resulting fragment into pcDNA5-FRT-TRP cut with

the same restriction enzymes. The resulting plasmid (TRP-LALA) was transfected into 293ATOA cells along with Flp recombinase (pOG44). Stably transfected cells were selected for at least two weeks in 100µg/µL hygromycin.

The MS2-only construct (pcNMS2-NLS-Fl) was previously described (Sahin et al., 2010). Using cDNA, we amplified full length PAPOLA (encoding PAPα) with forward and reverse primers containing BamHI and XbaI restriction sites, respectively. Following digestion, the PAPOLA gene was cloned into pnNMS2-NLS-Fl cut with the restriction enzymes. To make PAN-6MS2 we first inserted a BglII site 57 nt upstream of the cleavage and polyadenylation site of TRP-PANΔENE (TRP-PANΔ79 BglII-1020). Next, we PCR amplified the six MS2 sites using TetRP-PANΔENE-6MS2 (Stubbs et al., 2012) as a template with primers containing BglII restriction sites. After digestion with BglII, we cloned the resulting fragment into PAN-BglII-1020 cut with the same enzyme.

Data analysis

Venn diagrams were generated using Biovenn (Hulsen et al., 2008). Nuclear enrichment (NE) scores were found for the 13,114 genes with RPKM > 0.5 in the control datasets. NE scores were calculated by dividing $RPKM_{nuclear}$ by $RPKM_{total}$, with scores ranging from 0.45 to 2.00. We placed each gene into one of 32 bins based on its NE score. In Figure 21B, genes more nuclear than average are colored increasingly dark shades of blue and genes more cytoplasmic than average are colored increasingly dark shades of red. Heat maps were generated using the GENE-E software (http://www.broadinstitute.org/cancer/software/GENE-E/index.html).

RT-qPCR

RNA was harvested using TRI Reagent according to the manufacturer's protocol. Following extraction, RNA was treated with DNase for one hour at 37°C to remove genomic DNA contamination prior to qPCR analysis. A typical reaction consisted of 2 µg RNA, 40mM Tris 8.0, 10mM $MgSO_4$, 1mM $CaCl_2$, 40U RNasIN (Promega), and 60U of DNase (Promega) in a 30µL reaction volume. cDNA was made with random hexamers using standard protocols. qPCR was performed using iTaq Universal SYBR Green Supermix (Biorad).

Transcription pulse-chase assays

Transcription pulse-chase assays were performed as previously described (Bresson and Conrad, 2013). Standard northern blotting techniques were used to probe for PAN and 7SK.

Biotinylation and streptavidin selection

The biotinylation and streptavidin selection protocol are based off (Dolken et al., 2008; Zeiner et al., 2008) with some modifications. RNA biotinylation was performed using EZ-Link Biotin-HPDP (Pierce). The biotinylation reaction was carried out in a 200µL reaction mixture consisting of 40µg RNA, 20mM NaOAc pH=5.2, 1mM EDTA, 0.1% SDS, 0.2mg/mL Biotin-HPDP, and 50% N,N-dimethylformamide (DMF) for 3 hours at 25°C. Unconjugated biotin-HPDP was removed with three sequential chloroform extractions. After extraction of the aqueous phase, 20µL (10% v/v) of 10M NH_4OAc was added to each tube, and the RNA was precipitated in 70% ethanol.

Streptavidin selection was carried out using magnetic Streptavidin T1 beads (Invitrogen), and a Dynal bead magnet (Invitrogen). Prior to use, the beads were washed three times in a 0.1X MPG solution (1X MPG was 1M NaCl, 10mM EDTA, and 100mM Tris 7.5) supplemented with 0.1% igepal. After the final wash, the beads were resuspended in a 1mL solution consisting of

0.1X MPG supplemented with 0.1% igepal, 0.1µg/µL poly(A) (Sigma-Aldrich), 0.1µg/µL ssDNA, 0.1 µg/µL cRNA, and 0.1% SDS, and blocked for one hour. RNA was precipitated, resuspended in a volume of 63µL water, and denatured at 65°C for 5 minutes. Next, RNA was incubated together with beads for one hour while nutating at room temperature. Non-biotinylated RNAs were removed with a series of 400µL washes: 0.1X MPG, 0.1X MPG at 55°C, 0.1X MPG, 1X MPG, 1X MPG, 0.1X MPG, 1X MPG without NaCl, 0.1X MPG. With the exception of the second wash, each solution included 0.1% igepal. Biotinylated RNAs were eluted twice for 5 minutes each in a 200µL solution of 0.1X MPG containing 5% β-mercaptoethanol, a reducing agent which cleaves the disulfide bond between the 4SU and biotin. The first elution step was at 25°C and the second was at 65°C. The two eluted fractions were combined and extracted with PCA once and chloroform twice. After extraction, 40µL (10% v/v) of 10M NH_4OAc was added to each tube, and the RNA was precipitated in 70% ethanol. RNA was reverse transcribed prior to RT-qPCR analysis.

Nuclear Run-on assay

For each knockdown condition, we prepared two fully confluent 10 cm plates of 293ATOA cells (~$2x10^7$ cells/knockdown). For the siControl knockdown we prepared two additional 10 cm plates for a no 4-thioUTP control. Cells were trypsinized and the cell pellet was lysed in 1mL of hypotonic lysis buffer (HLB) (10mM Tris 7.5, 10mM NaCl, 2.5mM $MgCl_2$, 1mM DTT, and 0.5% igepal) for five minutes on ice. The lysate was underlayed with 1mL of HLB supplemented with 25% sucrose, and centrifuged at 600g/5'/4°C in order to isolate nuclei. While on ice, the pelleted nuclei (approximately 60µL in volume) were resuspended in 60µL of 2X transcription buffer (20mM Tris 8.0, 180mM KCl, 10mM $MgCl_2$, 50% glycerol, 5mM DTT, 40U RNasIN (Promega), and 250µM each of rATP, rCTP, and rGTP) supplemented with 1µM

of either 4-thioUTP or rUTP (for the no-4-thioUTP control). 3 µL of a 20% sarkosyl solution was added to each sample with gentle mixing in order to prevent new transcription initiation events. The run-on step was performed for 5 minutes at 30°C before RNA was harvested with TRI reagent. Purified RNA was treated with DNase as described above, with the exception that the reaction was halted with the addition of 25mM EDTA and 15mM EGTA. Following PCA extraction, the RNA was precipitated with ethanol and resuspended to a final concentration of 1mg/mL. 40µL of RNA from each sample was added to 10µL of 1M NaOH, and incubated on ice for precisely 4 minutes in order to hydrolyze the RNA. This step was used in order to better resolve the positions of elongating polymerases. The reaction was stopped with the addition of 61µL of neutralizing solution (833mM Tris 6.8, 500mM NaOAc, and 1µL of glycoblue (Ambion)). RNA was PCA extracted, and biotinylated and selected using magnetic streptavidin beads as described above. Selected RNA was reverse transcribed prior to RT-qPCR analysis.

4SU pulse chase

Following knockdown, cells were treated with 2µM of 4SU for one hour. Afterwards, cells were washed twice with phosphate buffered saline (PBS) containing calcium and magnesium (Sigma-Aldrich), and grown in media lacking 4SU for an additional hour. After the one hour washout step, we collected t=0, 30', 60', and 120' time points. 40µg of RNA was used as input for a biotinylation and streptavidin selection as described above. Selected RNA was reverse transcribed prior to RT-qPCR analysis. β-actin was used as a loading control for qPCR analysis.

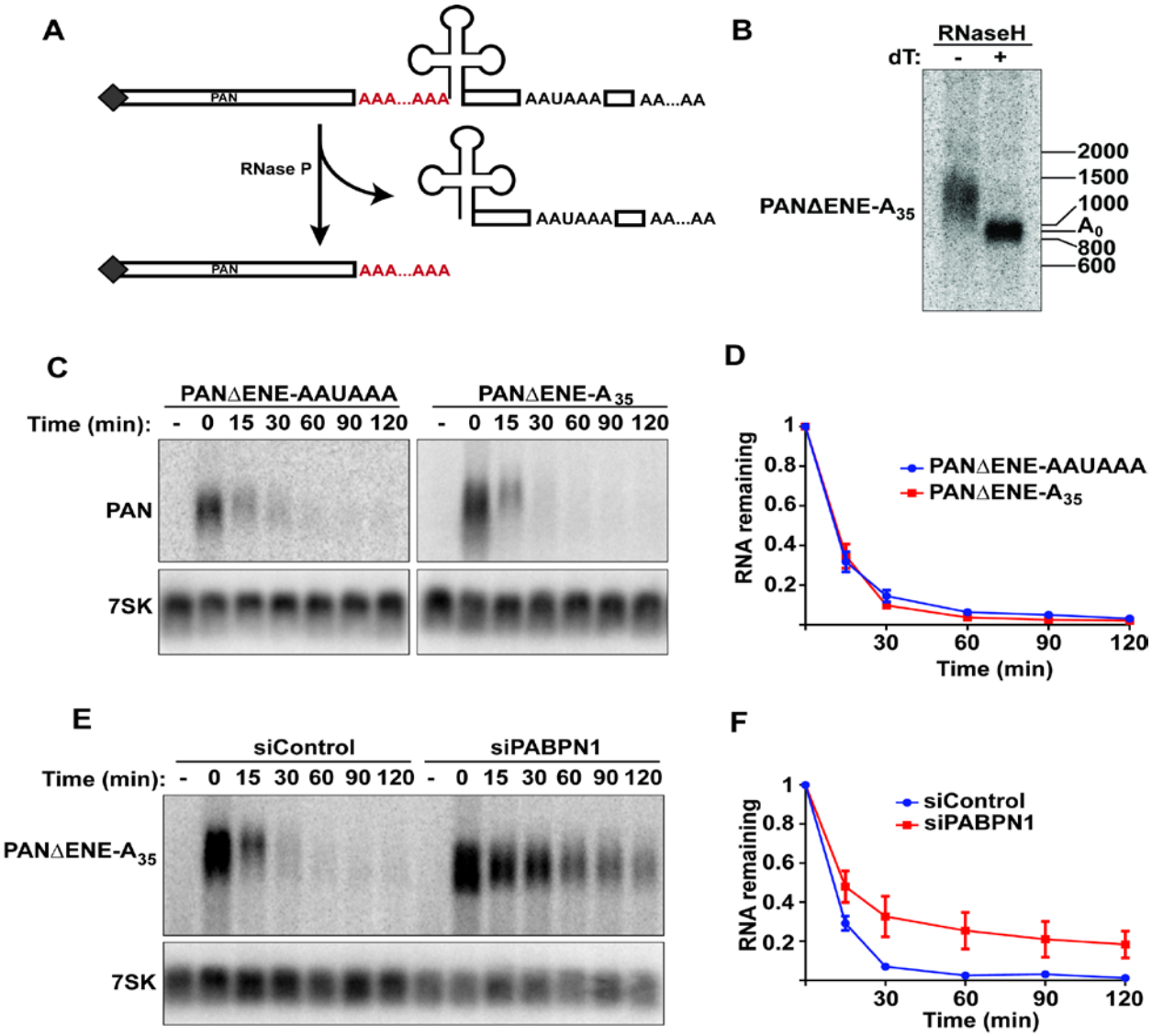

Figure 16. **CPSF-mediated polyadenylation is not required for decay. (A)** Schematic showing how PANΔENE was posttranscriptionally cleaved by RNase P. The 35 nt genomically encoded poly(A) tail is shown in red. The mascRNA sequence is represented by the clover-leaf structure. **(B)** Poly(A) tail length analysis of PAN-A_{35}. RNA was harvested and treated with RNase H in the presence or absence of oligo(dT) in order to degrade the poly(A) tail. **(C)** Representative transcription pulse-chase assay comparing PAN derived from a normal cleavage and polyadenylation reaction (PANΔENE-AAUAAA) or PAN derived from RNase P-mediated cleavage (PANΔENE-A_{35}). The "-" samples were harvested prior to the two hour transcription pulse. 7SK (bottom) was used as a loading control. **(D)** Quantification of the results in **(C)** *(n=3)*. Error bars show the standard deviation of the mean. **(E and F)** Representative transcription pulse chase and quantification *(n=3)* of PANΔENE-A_{35} following treatment with the indicated siRNAs.

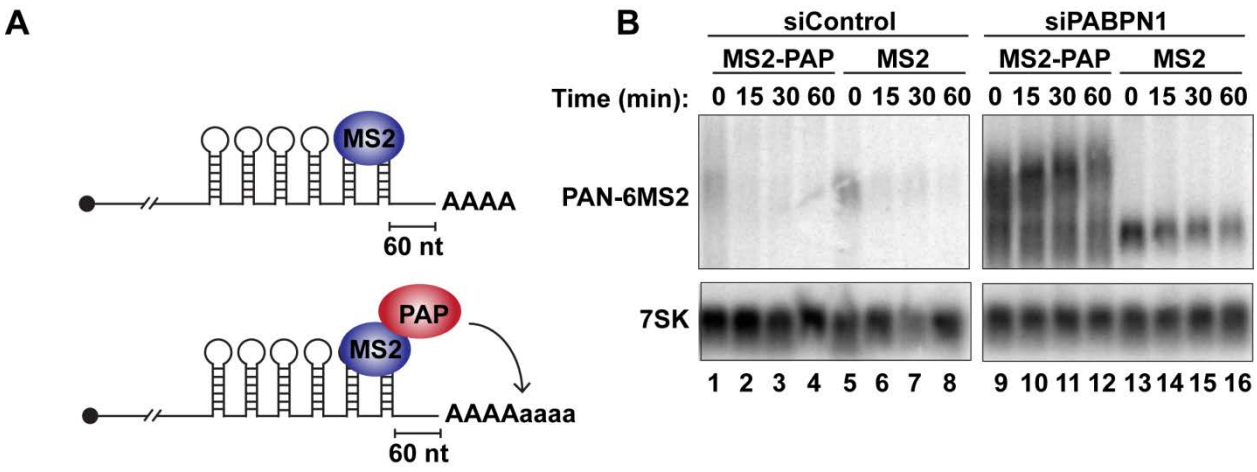

Figure 17. Polyadenylation is not sufficient for PABPN1-mediated decay. (A) Diagram illustrating the PAP-tethering approach. Six MS2 binding sites were inserted 60 nt upstream of the poly(A) tail of PANΔENE. Cells were cotransfected with either MS2 binding protein alone (top), or MS2-binding protein fused to PAP (bottom). **(B)** Transcription pulse-chase analysis of PANΔENE containing six MS2 binding sites. Cells were treated with either a control siRNA (left) or siRNAs targeting PABPN1 (right). Following knockdown cells were cotransfected with PANΔENE-6MS2 and either MS2-PAP or MS2 alone.

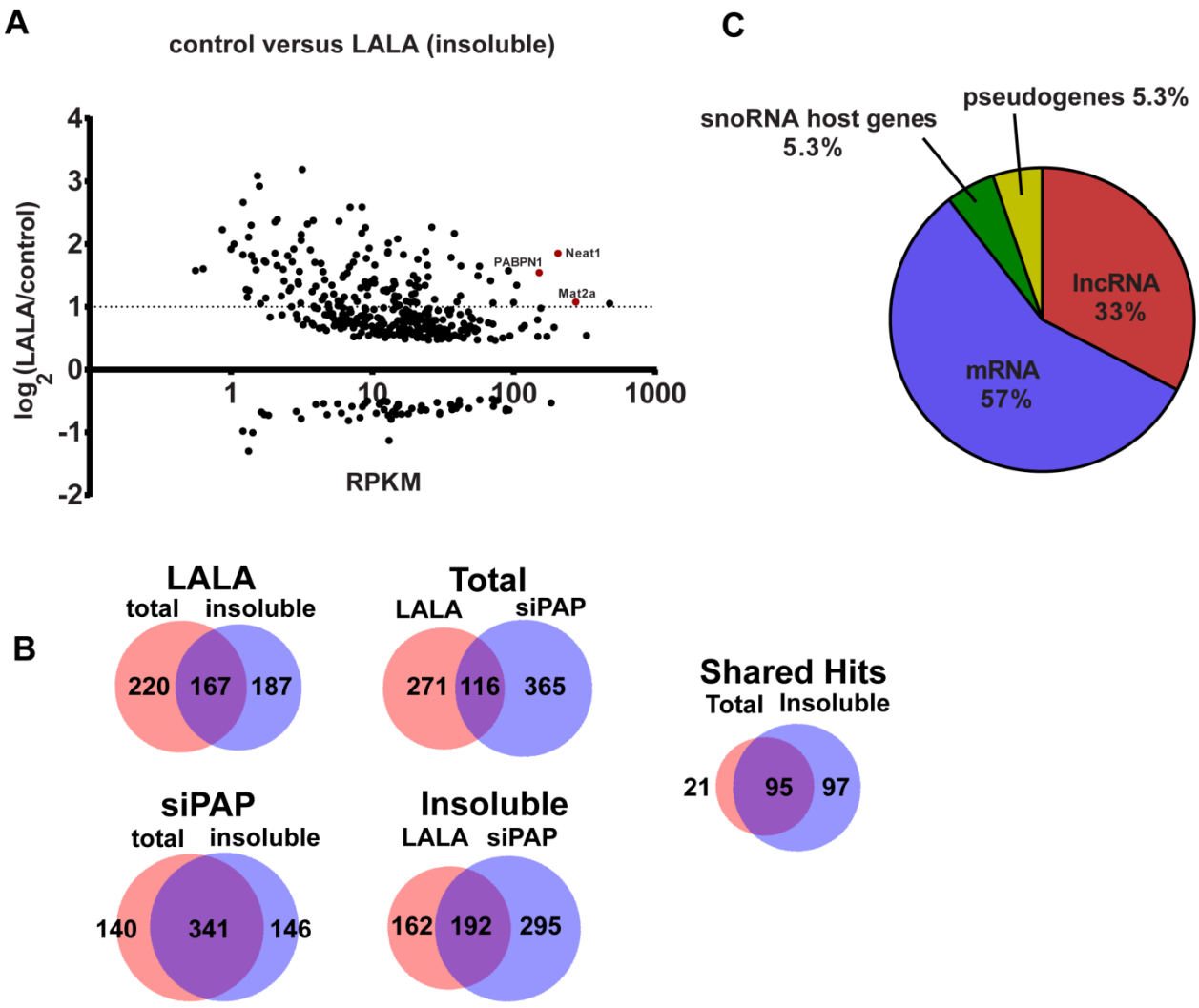

Figure 18. Global analysis of PPD targets. (A) Graph of differentially expressed genes in the nuclear enriched fractions from control cells versus cells stably expressing PABPN1-LALA. The $\log_2(RPKM_{LALA} / RPKM_{control})$ is plotted on the y-axis, while expression level (measured by RPKM) is shown on the x-axis. **(B)** Venn diagrams showing overlap in differentially expressed genes ($p<.05$ and FDR<5%) between different conditions. **(C)** Pie chart of the gene annotations for the 95 DEGs identified in all four datasets.

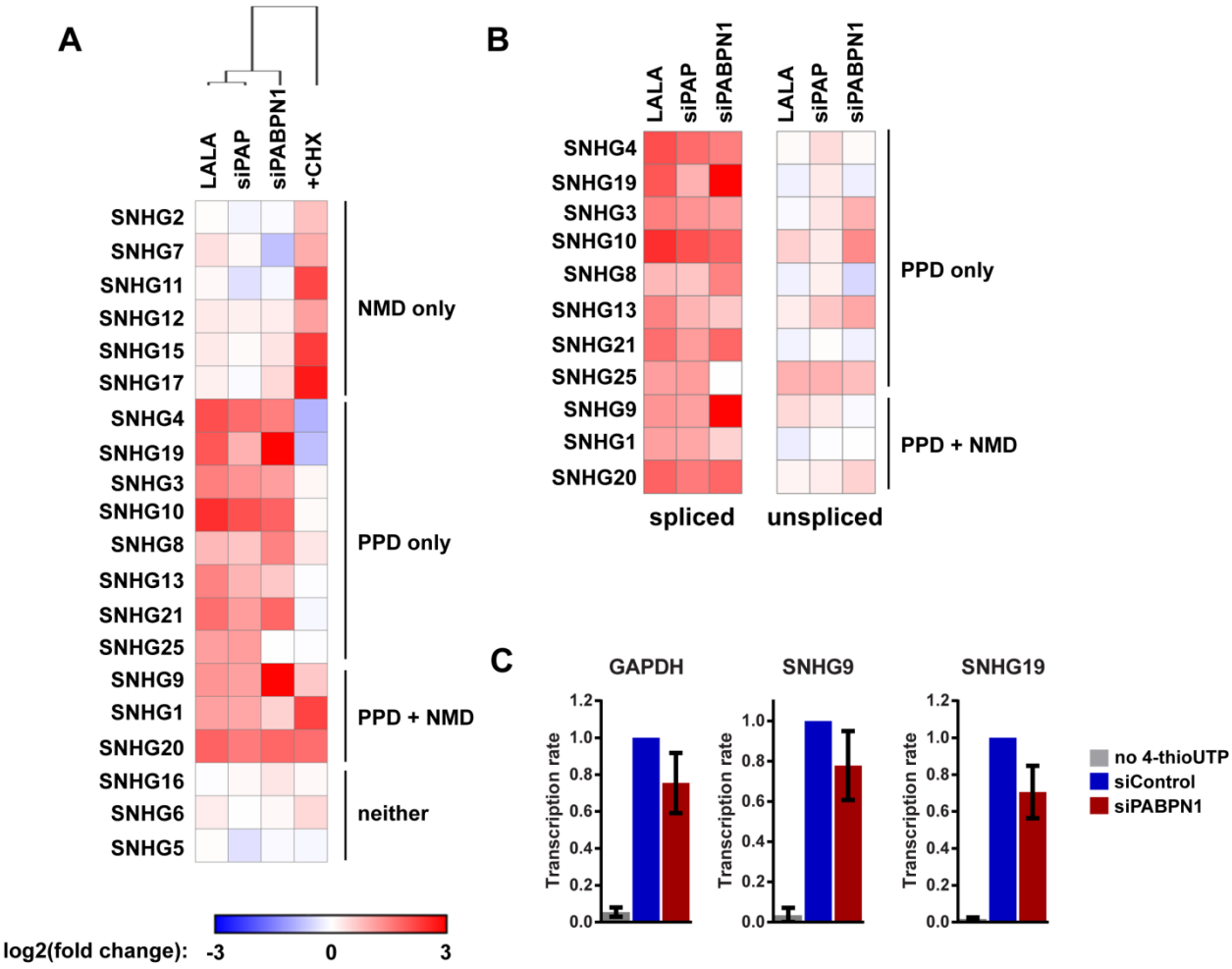

Figure 19. Spliced snoRNA host genes are targeted for decay by PPD. (A) Heat map showing the changes in spliced SNHG levels following LALA expression, PAP knockdown, PABPN1 knockdown, or cycloheximide treatment. The color of each square reflects the $\log_2$ value of the average change in transcript levels as measured by RT-qPCR *(n=3)*. Red shades mean the transcript is increased relative to the control, while blue shades indicate the transcript is decreased relative to the control. The red-blue scale is shown on the bottom. β-actin was used as a loading control for LALA, siPAP, and siPABPN1, and 7SK was used as a loading control for the cycloheximide experiment. **(B)** Same as in **(A)**, but showing the relative changes in both spliced and unspliced transcripts. **(C)** Results of a nuclear run-on following PABPN1 depletion (see text for details). Error bars show the average and standard deviation of three independent experiments.

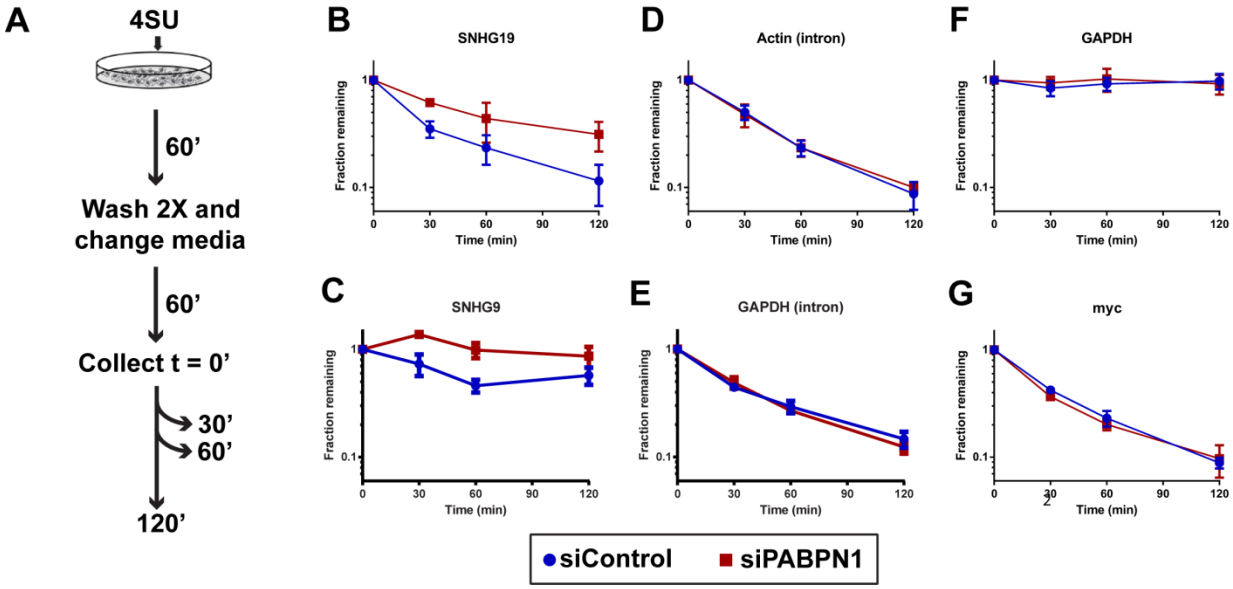

Figure 20. snoRNA host genes are stabilized in the absence of PABPN1. (A) Outline of the 4SU pulse-chase experiment (see text for details). **(B** through **G)** The results of a 4SU pulse chase from cells treated with either a control siRNA (blue) or siRNAs targeting PABPN1 (red) *(n=3)*. β-actin was used as a loading control.

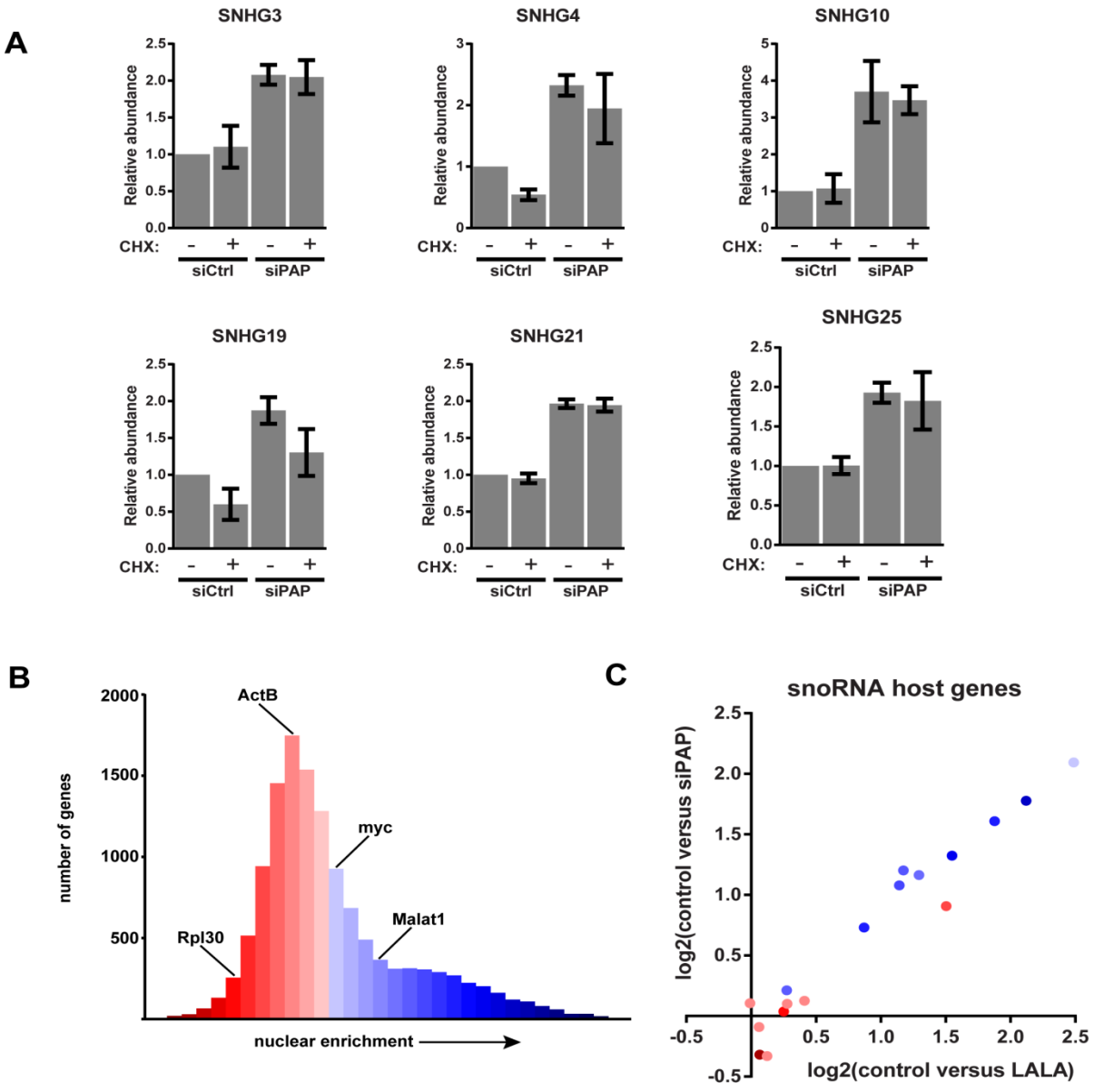

Figure 21. Nuclear retained snoRNA host genes are preferentially targeted by PPD. (A) RT-qPCR analysis showing the relative levels of six SNHGs following PAP depletion and the addition of cycloheximide. **(B)** Genome wide analysis showing the relative nuclear enrichment for each gene in the control dataset with RPKM > 0.5. Nuclear-enrichment was calculated by dividing $RPKM_{nuclear}$ by $RPKM_{total}$ for each transcript. Transcripts which are more cytoplasmic than average are shown in increasingly dark shades of red, while transcripts which are more nuclear than average are shown in increasingly dark shades of blue. The nuclear enrichment values of several well-known nuclear and cytoplasmic genes are shown for comparison. **(C)** Plot showing the nuclear enrichment of the SNHGs examined in this study. SNHGs are plotted by the fold enrichment values determined in Figure 19A. Values represent the average of three replicates. SNHGs are colored by their nuclear enrichment score determined in **(B)**.

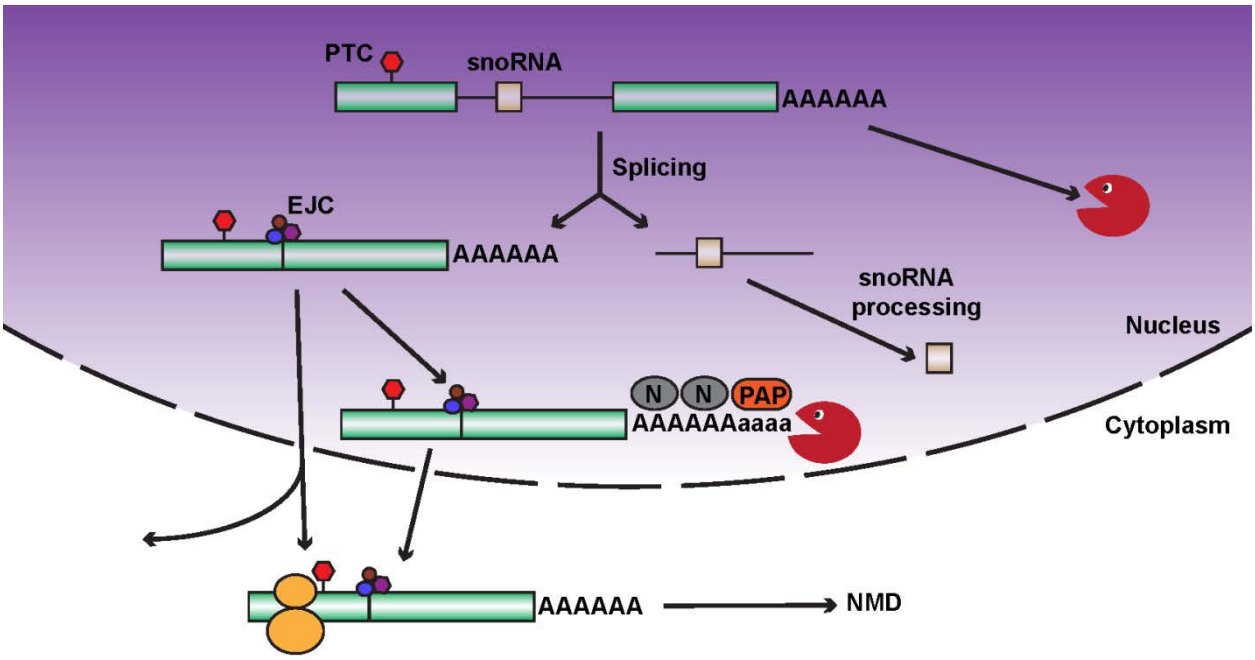

Figure 22. Model of snoRNA host gene decay. Symbols: PABPN1 (grey oval inscribed with "N"); Exonuclease (red pacman); Exon junction complex (EJC, trimeric complex sitting on the splice junction); Premature termination codon (PTC, red hexagon); ribosome (double orange ovals).

Host gene	snoRNA(s) encoded	Notes
SNHG1	SNORD25, SNORD26 SNORD27, SNORD28, SNORD29, SNORD30, SNORD31, SNORD22	
SNHG2	SNORD74, SNORD75, SNORD76, SNORD77, SNORD44, SNORD78, SNORD79, SNORD80, SNORD47, SNORD81	
SNHG3	SNORA73A, RNU105A	
SNHG4	SNORA74A	
SNHG5	SNORD50B, SNORD50A	
SNHG6	SNORD87	
SNHG7	SNORA43, SNORA17	
SNHG8	SNORA24	
SNHG9	SNORA78	
SNHG10	SCARNA13	
SNHG11	SNORA71E, SNORA60	
SNHG12	SNORD99, SNORA16A, SNORA61, SNORA44	
SNHG13	SNORA26	
SNHG14		Not in Refseq
SNHG15	SNORA9	
SNHG16	SNORD1A, SNORD1B, SNORD1C	
SNHG17	SNORA71A, SNORA71B, SNORA71C, SNORA71D	
SNHG18	SNORD123	Not expressed in 293ATOA cells.
SNHG19	SNORD60	
SNHG20	SCARNA16	
SNHG21	SCARNA15	
SNHG22	SCARNA17	Very lowly expressed and unclear annotation. Excluded from analysis.
SNHG23	SNORD113-1, SNORD113-2, SNORD113-3, SNORD113-4, SNORD113-5, SNORD113-6, SNORD113-7, SNORD113-8, SNORD113-9, SNORD114-1, SNORD114-2, SNORD114-3, SNORD114-4, SNORD114-5, SNORD114-6	Not expressed in 293ATOA cells.
SNHG24	SNORD114-11, SNORD114-12, SNORD114-13, SNORD114-14, SNORD114-15, SNORD114-16, SNORD114-17, SNORD114-18, SNORD114-19, SNORD114-20, SNORD114-21, SNORD114-22, SNORD114-23, SNORD114-24, SNORD114-25	Not expressed in 293ATOA cells.

Table 3. List of noncoding snoRNA host genes.

Chapter 4: Conclusions

Despite significant progress in uncovering the mechanisms of nuclear RNA decay in yeast, these processes remained relatively unexplored in higher eukaryotes. In human cells, nuclear RNAs which fail to undergo correct processing are rapidly degraded, suggesting the existence of an active nuclear quality control pathway. In this project, we characterized a novel nuclear quality control pathway which targets intronless RNAs. We identified the components of the decay pathway (PABPN1, PAP, and the nuclear exosome), and showed that PAN and intronless β-globin are both targeted by these factors (Chapter 2). Importantly, PABPN1-PAP decay also regulates the steady state levels of endogenous RNAs, including noncoding snoRNA host genes (Chapter 3). Taken together, this work contributes to a more complete understanding of the mechanisms involved in nuclear gene expression.

Mechanism of PPD

For all decay targets which we have examined in detail (PANΔENE, intronless β-globin, and noncoding SNHGs), active polyadenylation is a requirement for degradation. How might polyadenylation be linked to RNA decay? We propose that PABPN1 and PAP may provide the exosome a suitable binding site by stimulating the polyadenylation of a "naked" poly(A) tail (Figure 15). Recent biochemical and structural studies showed that ~30 nt of naked RNA is required to bind the exosome central channel in order to productively engage DIS3 for processive exosome degradation (Bonneau et al., 2009; Makino et al., 2013; Wasmuth and Lima, 2012). Poly(A) tail extension may be required to transiently produce unbound poly(A) 3´ ends sufficiently long enough to extend through the exosome channel. In this model, poly(A) tail

extension leads to a kinetic competition between additional PAP binding and exosome binding. When the exosome wins the competition, its processive activity will lead to rapid RNA destruction. If PAP re-binds, the process repeats itself. This model suggests that decay is inhibited when polyadenylation is highly processive. Under these conditions, PAP is tightly tethered 3´ end of the RNA, rendering it inaccessible to the exosome complex. Consistent with this model, CPSF, which mediates the highly processive phase of polyadenylation, is not required for decay (Figure 16). Conversely, when polyadenylation is only mildly processive, as is the case for poly(A) tails greater than ~200 nt (Eckmann et al., 2011), repeated binding and dissociation by PAP may allow the exosome transient access to the 3´ end, leading to degradation of the RNA.

Surprisingly, polyadenylation is not sufficient to rescue decay in the absence of PABPN1 (Figure 17). This observation suggests that PABPN1 has an additional role in decay apart from stimulating polyadenylation. We hypothesize that PABPN1 may further contribute to decay by direct recruitment of the exosome. The exosome has relatively weak affinity for poly(A) RNA (Wasmuth and Lima, 2012), and may require “bridging” factors in order to efficiently degrade RNA. In fact, *S. pombe* Pab2 appears to function in this fashion (see discussion section in Chapter 2). Notably, PABPN1 has been reported to physically interact with RRP6 and RRP40 (Beaulieu et al., 2012), further supporting this mechanism. This hypothesis remains speculative and mutagenesis experiments are needed to confirm that the physical interaction between PABPN1 and the exosome is indeed important for decay.

Noncoding snoRNA host genes are degraded by PPD

In Chapter 2, we identified the viral transcript PAN RNA as a target of PPD. Like many other herpes virus RNAs, PAN is intronless, and thus potentially subject to host quality control pathways which target unspliced transcripts. To counteract host decay pathways, PAN RNA has evolved the ENE, a 79 nt element which interacts in cis with the poly(A) tail. The ENE sequesters the 3' end of the transcript, protecting it from PPD. Interestingly, other viral transcripts and at least two human intronless RNAs also contain ENE-like elements, suggesting this may be a general mechanism for evading nuclear quality control pathways (Brown et al., 2012; Tycowski et al., 2012; Wilusz et al., 2012). Some naturally intronless mRNAs in humans have evolved cytoplasmic accumulation regions (CARs), which presumably allow CAR-containing transcripts to escape nuclear decay through more efficient export (Lei et al., 2011).

SNHGs present an interesting contrast to PAN RNA. Because snoRNAs are only encoded within introns, snoRNA host genes must undergo splicing to allow for proper snoRNA biogenesis. Thus, it is somewhat surprising that a significant number of spliced SNHGs would be retained in the nucleus and subject to a quality control pathway which primarily targets intronless RNAs. We hypothesize that SNHGs are degraded within the nucleus in order to avoid deleterious consequences arising from inappropriate translation in the cytoplasm. Noncoding SNHGs lack an open reading frame, and translation could lead to the production of short, non-functional peptides with potentially dominant negative phenotypes. Normally, RNAs with aberrant stop codons are degraded by NMD, and we found that 9 SNHGs are degraded in this manner. However, no decay pathway is completely efficient, and some RNAs may escape degradation. In addition, snoRNA host genes tend to be highly transcribed, and could represent a substantial burden for the NMD machinery. Consequently, SNHGs appear to have evolved

greater susceptibility to PPD, allowing transcripts to be degraded before they reach the cytoplasm.

In addition to snoRNA host genes, our RNAseq analysis uncovered several other classes of genes targeted by PPD, including promoter antisense lncRNAs, intergenic lncRNAs, noncoding miRNA host genes and mRNAs. Moreover, many upregulated RNAs were completely unannotated, suggesting that our analysis is an underestimate of the true extent of PPD. Future studies will focus on PPD's role in the regulation of these still uncharacterized decay targets.

References

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. The EMBO journal *18*, 5399-5410.

Apponi, L.H., Kelly, S.M., Harreman, M.T., Lehner, A.N., Corbett, A.H., and Valentini, S.R. (2007). An interaction between two RNA binding proteins, Nab2 and Pub1, links mRNA processing/export and mRNA stability. Molecular and cellular biology *27*, 6569-6579.

Apponi, L.H., Leung, S.W., Williams, K.R., Valentini, S.R., Corbett, A.H., and Pavlath, G.K. (2010). Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis. Human molecular genetics *19*, 1058-1065.

Ares, M., Jr., Grate, L., and Pauling, M.H. (1999). A handful of intron-containing genes produces the lion's share of yeast mRNA. RNA *5*, 1138-1139.

Askarian-Amiri, M.E., Crawford, J., French, J.D., Smart, C.E., Smith, M.A., Clark, M.B., Ru, K., Mercer, T.R., Thompson, E.R., Lakhani, S.R., *et al.* (2011). SNORD-host RNA Zfas1 is a regulator of mammary development and a potential marker for breast cancer. RNA *17*, 878-891.

Beaulieu, Y.B., Kleinman, C.L., Landry-Voyer, A.M., Majewski, J., and Bachand, F. (2012). Polyadenylation-dependent control of long noncoding RNA expression by the poly(A)-binding protein nuclear 1. PLoS genetics *8*, e1003078.

Benoit, B., Mitou, G., Chartier, A., Temme, C., Zaessinger, S., Wahle, E., Busseau, I., and Simonelig, M. (2005). An essential cytoplasmic function for the nuclear poly(A) binding protein, PABP2, in poly(A) tail length control and early development in Drosophila. Developmental cell *9*, 511-522.

Benoit, B., Nemeth, A., Aulner, N., Kuhn, U., Simonelig, M., Wahle, E., and Bourbon, H.M. (1999). The Drosophila poly(A)-binding protein II is ubiquitous throughout Drosophila development and has the same function in mRNA polyadenylation as its bovine homolog in vitro. Nucleic acids research *27*, 3771-3778.

Bhattacharjee, R.B., and Bag, J. (2012). Depletion of nuclear poly(A) binding protein PABPN1 produces a compensatory response by cytoplasmic PABP4 and PABP5 in cultured human cells. PloS one *7*, e53036.

Bonneau, F., Basquin, J., Ebert, J., Lorentzen, E., and Conti, E. (2009). The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. Cell *139*, 547-559.

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

Brown, J.A., Valenstein, M.L., Yario, T.A., Tycowski, K.T., and Steitz, J.A. (2012). Formation of triple-helical structures by the 3'-end sequences of MALAT1 and MENbeta noncoding RNAs. Proceedings of the National Academy of Sciences of the United States of America *109*, 19202-19207.

Buchman, A.R., and Berg, P. (1988). Comparison of intron-dependent and intron-independent gene expression. Molecular and cellular biology *8*, 4395-4405.

Callahan, K.P., and Butler, J.S. (2010). TRAMP complex enhances RNA degradation by the nuclear exosome component Rrp6. The Journal of biological chemistry *285*, 3540-3547.

Callis, J., Fromm, M., and Walbot, V. (1987). Introns increase gene expression in cultured maize cells. Genes & development *1*, 1183-1200.

Chang, Y.F., Imam, J.S., and Wilkinson, M.F. (2007). The nonsense-mediated decay RNA surveillance pathway. Annual review of biochemistry *76*, 51-74.

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.

Chen, I.H., Sciabica, K.S., and Sandri-Goldin, R.M. (2002). ICP27 interacts with the RNA export factor Aly/REF to direct herpes simplex virus type 1 intronless mRNAs to the TAP export pathway. Journal of virology *76*, 12877-12889.

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.

Church, G.M., and Gilbert, W. (1984). Genomic sequencing. Proceedings of the National Academy of Sciences of the United States of America *81*, 1991-1995.

Conrad, N.K., Mili, S., Marshall, E.L., Shu, M.D., and Steitz, J.A. (2006). Identification of a rapid mammalian deadenylation-dependent decay pathway and its inhibition by a viral RNA element. Molecular cell *24*, 943-953.

Conrad, N.K., Shu, M.D., Uyhazi, K.E., and Steitz, J.A. (2007). Mutational analysis of a viral RNA element that counteracts rapid RNA decay by interaction with the polyadenylate tail. Proceedings of the National Academy of Sciences of the United States of America *104*, 10412-10417.

Conrad, N.K., and Steitz, J.A. (2005). A Kaposi's sarcoma virus RNA element that increases the nuclear abundance of intronless transcripts. The EMBO journal *24*, 1831-1841.

Davidson, L., Kerr, A., and West, S. (2012). Co-transcriptional degradation of aberrant pre-mRNA by Xrn2. The EMBO journal *31*, 2566-2578.

de Klerk, E., Venema, A., Anvar, S.Y., Goeman, J.J., Hu, O., Trollet, C., Dickson, G., den Dunnen, J.T., van der Maarel, S.M., Raz, V., *et al.* (2012). Poly(A) binding protein nuclear 1 levels affect alternative polyadenylation. Nucleic acids research *40*, 9089-9101.

Dheur, S., Nykamp, K.R., Viphakone, N., Swanson, M.S., and Minvielle-Sebastia, L. (2005). Yeast mRNA Poly(A) tail length control can be reconstituted in vitro in the absence of Pab1p-dependent Poly(A) nuclease activity. The Journal of biological chemistry *280*, 24532-24538.

Dolken, L., Ruzsics, Z., Radle, B., Friedel, C.C., Zimmer, R., Mages, J., Hoffmann, R., Dickinson, P., Forster, T., Ghazal, P., *et al.* (2008). High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. RNA *14*, 1959-1972.

Doma, M.K., and Parker, R. (2007). RNA quality control in eukaryotes. Cell *131*, 660-668.

Eckmann, C.R., Rammelt, C., and Wahle, E. (2011). Control of poly(A) tail length. Wiley interdisciplinary reviews RNA *2*, 348-361.

Fasken, M.B., and Corbett, A.H. (2009). Mechanisms of nuclear mRNA quality control. RNA biology *6*, 237-241.

Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W., and Luhrmann, R. (1995). The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. Cell *82*, 475-483.

Goebels, C., Thonn, A., Gonzalez-Hilarion, S., Rolland, O., Moyrand, F., Beilharz, T.H., and Janbon, G. (2013). Introns regulate gene expression in Cryptococcus neoformans in a Pab2p dependent pathway. PLoS genetics *9*, e1003686.

Grzechnik, P., and Kufel, J. (2008). Polyadenylation linked to transcription termination directs the processing of snoRNA precursors in yeast. Molecular cell *32*, 247-258.

Guang, S., Felthauser, A.M., and Mertz, J.E. (2005). Binding of hnRNP L to the pre-mRNA processing enhancer of the herpes simplex virus thymidine kinase gene enhances both polyadenylation and nucleocytoplasmic export of intronless mRNAs. Molecular and cellular biology *25*, 6303-6313.

Hautbergue, G.M., Hung, M.L., Walsh, M.J., Snijders, A.P., Chang, C.T., Jones, R., Ponting, C.P., Dickman, M.J., and Wilson, S.A. (2009). UIF, a New mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. Current biology : CB *19*, 1918-1924.

Hector, R.E., Nykamp, K.R., Dheur, S., Anderson, J.T., Non, P.J., Urbinati, C.R., Wilson, S.M., Minvielle-Sebastia, L., and Swanson, M.S. (2002). Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. The EMBO journal *21*, 1800-1810.

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.

Hilleren, P., and Parker, R. (2001). Defects in the mRNA export factors Rat7p, Gle1p, Mex67p, and Rat8p cause hyperadenylation during 3'-end formation of nascent transcripts. RNA *7*, 753-764.

Holbein, S., Wengi, A., Decourty, L., Freimoser, F.M., Jacquier, A., and Dichtl, B. (2009). Cordycepin interferes with 3' end formation in yeast independently of its potential to terminate RNA chain elongation. RNA *15*, 837-849.

Hosoda, N., Lejeune, F., and Maquat, L.E. (2006). Evidence that poly(A) binding protein C1 binds nuclear pre-mRNA poly(A) tails. Molecular and cellular biology *26*, 3085-3097.

Houseley, J., Kotovic, K., El Hage, A., and Tollervey, D. (2007). Trf4 targets ncRNAs from telomeric and rDNA spacer regions and functions in rDNA copy number control. The EMBO journal *26*, 4996-5006.

Hulsen, T., de Vlieg, J., and Alkema, W. (2008). BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC genomics *9*, 488.

Jao, C.Y., and Salic, A. (2008). Exploring RNA transcription and turnover in vivo by using click chemistry. Proceedings of the National Academy of Sciences of the United States of America *105*, 15779-15784.

Jenal, M., Elkon, R., Loayza-Puch, F., van Haaften, G., Kuhn, U., Menzies, F.M., Oude Vrielink, J.A., Bos, A.J., Drost, J., Rooijers, K., *et al.* (2012). The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. Cell *149*, 538-553.

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

Jia, H., Wang, X., Liu, F., Guenther, U.P., Srinivasan, S., Anderson, J.T., and Jankowsky, E. (2011). The RNA helicase Mtr4p modulates polyadenylation in the TRAMP complex. Cell *145*, 890-901.

Juge, F., Zaessinger, S., Temme, C., Wahle, E., and Simonelig, M. (2002). Control of poly(A) polymerase level is essential to cytoplasmic polyadenylation and early development in Drosophila. The EMBO journal *21*, 6603-6613.

Kadaba, S., Krueger, A., Trice, T., Krecic, A.M., Hinnebusch, A.G., and Anderson, J. (2004). Nuclear surveillance and degradation of hypomodified initiator tRNAMet in S. cerevisiae. Genes & development *18*, 1227-1240.

Kahvejian, A., Svitkin, Y.V., Sukarieh, R., M'Boutchou, M.N., and Sonenberg, N. (2005). Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. Genes & development *19*, 104-113.

Keller, W., Bienroth, S., Lang, K.M., and Christofori, G. (1991). Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. The EMBO journal *10*, 4241-4249.

Kelly, S.M., Pabit, S.A., Kitchen, C.M., Guo, P., Marfatia, K.A., Murphy, T.J., Corbett, A.H., and Berland, K.M. (2007). Recognition of polyadenosine RNA by zinc finger proteins. Proceedings of the National Academy of Sciences of the United States of America *104*, 12306-12311.

Kerwitz, Y., Kuhn, U., Lilie, H., Knoth, A., Scheuermann, T., Friedrich, H., Schwarz, E., and Wahle, E. (2003). Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. The EMBO journal *22*, 3705-3714.

Kino, T., Hurt, D.E., Ichijo, T., Nader, N., and Chrousos, G.P. (2010). Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. Science signaling *3*, ra8.

Kiss, D.L., and Andrulis, E.D. (2011). The exozyme model: a continuum of functionally distinct complexes. RNA *17*, 1-13.

Kiss, T., and Filipowicz, W. (1995). Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. Genes & development *9*, 1411-1424.

Kuhn, U., Gundel, M., Knoth, A., Kerwitz, Y., Rudel, S., and Wahle, E. (2009). Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. The Journal of biological chemistry *284*, 22803-22814.

Kuhn, U., Nemeth, A., Meyer, S., and Wahle, E. (2003). The RNA binding domains of the nuclear poly(A)-binding protein. The Journal of biological chemistry *278*, 16916-16925.

Kyriakopoulou, C.B., Nordvarg, H., and Virtanen, A. (2001). A novel nuclear human poly(A) polymerase (PAP), PAP gamma. The Journal of biological chemistry *276*, 33504-33511.

LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005). RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell *121*, 713-724.

Lee, Y.J., and Glaunsinger, B.A. (2009). Aberrant herpesvirus-induced polyadenylation correlates with cellular messenger RNA destruction. PLoS biology *7*, e1000107.

Lei, H., Dias, A.P., and Reed, R. (2011). Export and stability of naturally intronless mRNAs require specific coding region sequences and the TREX mRNA export complex. Proceedings of the National Academy of Sciences of the United States of America *108*, 17985-17990.

Lemay, J.F., D'Amours, A., Lemieux, C., Lackner, D.H., St-Sauveur, V.G., Bahler, J., and Bachand, F. (2010). The nuclear poly(A)-binding protein interacts with the exosome to promote synthesis of noncoding small nucleolar RNAs. Molecular cell *37*, 34-45.

Lemieux, C., Marguerat, S., Lafontaine, J., Barbezier, N., Bahler, J., and Bachand, F. (2011). A Pre-mRNA degradation pathway that selectively targets intron-containing genes requires the nuclear poly(A)-binding protein. Molecular cell *44*, 108-119.

Loflin, P.T., Chen, C.Y., Xu, N., and Shyu, A.B. (1999). Transcriptional pulsing approaches for analysis of mRNA turnover in mammalian cells. Methods *17*, 11-20.

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. Molecular cell *43*, 624-637.

Makino, D.L., Baumgartner, M., and Conti, E. (2013). Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. Nature *495*, 70-75.

Malim, M.H., Hauber, J., Le, S.Y., Maizel, J.V., and Cullen, B.R. (1989). The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature *338*, 254-257.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997). The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. Cell *91*, 457-466.

Mitton-Fry, R.M., DeGregorio, S.J., Wang, J., Steitz, T.A., and Steitz, J.A. (2010). Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. Science *330*, 1244-1247.

Mohanty, B.K., and Kushner, S.R. (2011). Bacterial/archaeal/organellar polyadenylation. Wiley interdisciplinary reviews RNA *2*, 256-276.

Moore, M.J., and Proudfoot, N.J. (2009). Pre-mRNA processing reaches back to transcription and ahead to translation. Cell *136*, 688-700.

Pak, C., Garshasbi, M., Kahrizi, K., Gross, C., Apponi, L.H., Noto, J.J., Kelly, S.M., Leung, S.W., Tzschach, A., Behjati, F., *et al.* (2011). Mutation of the conserved polyadenosine RNA binding protein, ZC3H14/dNab2, impairs neural function in Drosophila and humans.Proceedings of the National Academy of Sciences of the United States of America *108*, 12390-12395.

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nature genetics *40*, 1413-1415.
Perreault, A., Lemieux, C., and Bachand, F. (2007). Regulation of the nuclear poly(A)-binding protein by arginine methylation in fission yeast. The Journal of biological chemistry *282*, 7552-7562.

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.

Qu, X., Lykke-Andersen, S., Nasser, T., Saguez, C., Bertrand, E., Jensen, T.H., and Moore, C. (2009). Assembly of an export-competent mRNP is needed for efficient release of the 3'-end processing complex after polyadenylation. Molecular and cellular biology *29*, 5327-5338.

Rearick, D., Prakash, A., McSweeny, A., Shepard, S.S., Fedorova, L., and Fedorov, A. (2011). Critical association of ncRNA with introns. Nucleic acids research *39*, 2357-2366.

Rose, A.B. (2004). The effect of intron location on intron-mediated enhancement of gene expression in Arabidopsis. The Plant journal : for cell and molecular biology *40*, 744-751.

Roth, K.M., Byam, J., Fang, F., and Butler, J.S. (2009). Regulation of NAB2 mRNA 3'-end formation requires the core exosome and the Trf4p component of the TRAMP complex. RNA *15*, 1045-1058.

Roth, K.M., Wolf, M.K., Rossi, M., and Butler, J.S. (2005). The nuclear exosome contributes to autogenous control of NAB2 mRNA levels. Molecular and cellular biology *25*, 1577-1585.

Rougemaille, M., Gudipati, R.K., Olesen, J.R., Thomsen, R., Seraphin, B., Libri, D., and Jensen, T.H. (2007). Dissecting mechanisms of nuclear mRNA surveillance in THO/sub2 complex mutants. The EMBO journal *26*, 2317-2326.

Ruegsegger, U., Beyer, K., and Keller, W. (1996). Purification and characterization of human cleavage factor Im involved in the 3' end processing of messenger RNA precursors. The Journal of biological chemistry *271*, 6107-6113.

Sachs, A.B., Davis, R.W., and Kornberg, R.D. (1987). A single domain of yeast poly(A)-binding protein is necessary and sufficient for RNA binding and cell viability. Molecular and cellular biology *7*, 3268-3276.

Saguez, C., Schmid, M., Olesen, J.R., Ghazy, M.A., Qu, X., Poulsen, M.B., Nasser, T., Moore, C., and Jensen, T.H. (2008). Nuclear mRNA surveillance in THO/sub2 mutants is triggered by inefficient polyadenylation. Molecular cell *31*, 91-103.

Sahin, B.B., Patel, D., and Conrad, N.K. (2010). Kaposi's sarcoma-associated herpesvirus ORF57 protein binds and protects a nuclear noncoding RNA from cellular RNA decay pathways. PLoS pathogens *6*, e1000799.

San Paolo, S., Vanacova, S., Schenk, L., Scherrer, T., Blank, D., Keller, W., and Gerber, A.P. (2009). Distinct roles of non-canonical poly(A) polymerases in RNA metabolism. PLoS genetics *5*, e1000555.

Schmid, M., and Jensen, T.H. (2008a). The exosome: a multipurpose RNA-decay machine. Trends in biochemical sciences *33*, 501-510.

Schmid, M., and Jensen, T.H. (2008b). Quality control of mRNP in the nucleus. Chromosoma *117*, 419-429.

Schmid, M., Poulsen, M.B., Olszewski, P., Pelechano, V., Saguez, C., Gupta, I., Steinmetz, L.M., Moore, C., and Jensen, T.H. (2012). Rrp6p controls mRNA poly(A) tail length and its decoration with poly(A) binding proteins. Molecular cell *47*, 267-280.

Schmidt, K., and Butler, J.S. (2013). Nuclear RNA surveillance: role of TRAMP in controlling exosome specificity. Wiley interdisciplinary reviews RNA *4*, 217-231.

Schmidt, M.J., and Norbury, C.J. (2010). Polyadenylation and beyond: emerging roles for noncanonical poly(A) polymerases. Wiley interdisciplinary reviews RNA *1*, 142-151.

Schneider, C., and Tollervey, D. (2013). Threading the barrel of the RNA exosome. Trends in biochemical sciences *38*, 485-493.

Shabalina, S.A., Ogurtsov, A.Y., Spiridonov, A.N., Novichkov, P.S., Spiridonov, N.A., and Koonin, E.V. (2010). Distinct patterns of expression and evolution of intronless and intron-containing mammalian genes. Molecular biology and evolution *27*, 1745-1749.

Shcherbik, N., Wang, M., Lapik, Y.R., Srivastava, L., and Pestov, D.G. (2010). Polyadenylation and degradation of incomplete RNA polymerase I transcripts in mammalian cells. EMBO reports *11*, 106-111.

Slomovic, S., Laufer, D., Geiger, D., and Schuster, G. (2006). Polyadenylation of ribosomal RNA in human cells. Nucleic acids research *34*, 2966-2975.

Slomovic, S., and Schuster, G. (2011). Exonucleases and endonucleases involved in polyadenylation-assisted RNA decay. Wiley interdisciplinary reviews RNA *2*, 106-123.

Smith, C.M., and Steitz, J.A. (1998). Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5'-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. Molecular and cellular biology *18*, 6897-6909.

St-Andre, O., Lemieux, C., Perreault, A., Lackner, D.H., Bahler, J., and Bachand, F. (2010). Negative regulation of meiotic gene expression by the nuclear poly(a)-binding protein in fission yeast. The Journal of biological chemistry *285*, 27859-27868.

Stubbs, S.H., Hunter, O.V., Hoover, A., and Conrad, N.K. (2012). Viral factors reveal a role for REF/Aly in nuclear RNA stability. Molecular and cellular biology *32*, 1260-1270.

Tomecki, R., and Dziembowski, A. (2010). Novel endoribonucleases as central players in various pathways of eukaryotic RNA metabolism. RNA *16*, 1692-1724.

Topalian, S.L., Kaneko, S., Gonzales, M.I., Bond, G.L., Ward, Y., and Manley, J.L. (2001). Identification and functional characterization of neo-poly(A) polymerase, an RNA processing enzyme overexpressed in human tumors. Molecular and cellular biology *21*, 5614-5623.

Tycowski, K.T., Shu, M.D., Borah, S., Shi, M., and Steitz, J.A. (2012). Conservation of a triple-helix-forming RNA stability element in noncoding and genomic RNAs of diverse viruses. Cell reports *2*, 26-32.

Tycowski, K.T., Shu, M.D., and Steitz, J.A. (1996). A mammalian gene with introns instead of exons generating stable RNA products. Nature *379*, 464-466.

Valencia, P., Dias, A.P., and Reed, R. (2008). Splicing promotes rapid and efficient mRNA export in mammalian cells. Proceedings of the National Academy of Sciences of the United States of America *105*, 3386-3391.

van Hoof, A., Lennertz, P., and Parker, R. (2000). Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. Molecular and cellular biology *20*, 441-452.

Vanacova, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G., and Keller, W. (2005). A new yeast poly(A) polymerase complex involved in RNA quality control. PLoS biology *3*, e189.

Viphakone, N., Voisinet-Hakil, F., and Minvielle-Sebastia, L. (2008). Molecular dissection of mRNA poly(A) tail length control in yeast. Nucleic acids research *36*, 2418-2433.

Wahle, E. (1991). A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. Cell *66*, 759-768.

Wahle, E. (1995). Poly(A) tail length control is caused by termination of processive synthesis. The Journal of biological chemistry *270*, 2800-2808.

Wahle, E., Martin, G., Schiltz, E., and Keller, W. (1991). Isolation and expression of cDNA clones encoding mammalian poly(A) polymerase. The EMBO journal *10*, 4251-4257.

Wasmuth, E.V., Januszyk, K., and Lima, C.D. (2014). Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA. Nature *511*, 435-439.

Wasmuth, E.V., and Lima, C.D. (2012). Exo- and endoribonucleolytic activities of yeast cytoplasmic and nuclear RNA exosomes are dependent on the noncatalytic core and central channel. Molecular cell *48*, 133-144.

West, S., Gromak, N., Norbury, C.J., and Proudfoot, N.J. (2006). Adenylation and exosome-mediated degradation of cotranscriptionally cleaved pre-messenger RNA in human cells. Molecular cell *21*, 437-443.

Wilusz, J.E., Freier, S.M., and Spector, D.L. (2008). 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. Cell *135*, 919-932.

Wilusz, J.E., JnBaptiste, 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 & development *26*, 2392-2407.

Winstall, E., Sadowski, M., Kuhn, U., Wahle, E., and Sachs, A.B. (2000). The Saccharomyces cerevisiae RNA-binding protein Rbp29 functions in cytoplasmic mRNA metabolism. The Journal of biological chemistry *275*, 21817-21826.

Wlotzka, W., Kudla, G., Granneman, S., and Tollervey, D. (2011). The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. The EMBO journal *30*, 1790-1803.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.C., Dufour, M.E., Boulay, J., Regnault, B., Devaux, F., Namane, A., Seraphin, B., *et al.* (2005). Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell *121*, 725-737.

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. The EMBO journal *29*, 2173-2181.

Zeiner, G.M., Cleary, M.D., Fouts, A.E., Meiring, C.D., Mocarski, E.S., and Boothroyd, J.C. (2008). RNA analysis by biosynthetic tagging using 4-thiouracil and uracil phosphoribosyltransferase. Methods in molecular biology *419*, 135-146.

Zhao, C., and Hamilton, T. (2007). Introns regulate the rate of unstable mRNA decay. The Journal of biological chemistry *282*, 20230-20237.