Sequence Determinants for Nuclear Retention and Cytoplasmic Export of mRNAs and lncRNAs

Alexander F. Palazzo* and Eliza S. Lee

Department of Biochemistry,

University of Toronto,

Toronto, ON, Canada

*Corresponding Author

alex.palazzo@utoronto.ca

Abstract

Eukaryotes are divided into two major compartments: the nucleus where RNA is synthesized and processed, and the cytoplasm, where mRNA is translated into proteins. Although many different RNAs are made, only a subset is allowed access to the cytoplasm, primarily RNAs involved in protein synthesis (mRNA, tRNA and rRNA). In contrast, nuclear retained transcripts are mostly long non-coding RNAs (lncRNAs) whose role in cell physiology has been a source of much investigation in the past few years. In addition, it is likely that many non-functional RNAs, which arise by spurious transcription and misprocessing of functional RNAs, are also retained in the nucleus and degraded. In this review, the main sequence features that dictate whether any particular mRNA or lncRNA is a substrate for retention in the nucleus, or export to the cytoplasm, are discussed. Although nuclear export is promoted by RNA-splicing due to the fact that the spliceosome can help recruit export factors to the mature RNA, nuclear export does not require splicing. Indeed, most stable unspliced transcripts are well exported and associate with these same export factors in a splicing-independent manner. In contrast, nuclear retention is promoted by specialized *cis*-elements found in certain RNAs. This new understanding of the determinants of nuclear retention and cytoplasmic export provides a deeper understanding of how information flow is regulated in eukaryotic cells. Ultimately these processes promote the evolution of complexity in eukaryotes by shaping the genomic content through constructive neutral evolution.

Key words

mRNA nuclear export, mRNA nuclear retention, lncRNA nuclear retention, TREX, lncRNAs, transposable elements, RNA modification, splicing, polyadenylation, Constructive Neutral Evolution

1. Introduction

The distinguishing feature of eukaryotic cells is that they are divided into two compartments: the nucleus where pre-messenger RNAs (mRNAs) are made and processed, and the cytoplasm where mature mRNAs are translated into proteins (Martin and Koonin, 2006; Palazzo and Gregory, 2014). This is in contrast to prokaryotes, where mRNAs are made and translated at the same time in the same compartment. In eukaryotes, the temporal and spatial separation of mRNA synthesis from translation allows each newly made RNA to be subjected to extensive quality control before it ever encounters a ribosome (Palazzo and Akef, 2012). This quality control involves the nuclear retention and/or degradation of spurious transcripts, which are synthesized from intergenic DNA regions, and misprocessed RNAs, which result from errors in splicing or 3' cleavage. In the absence of this quality control, spurious transcripts and misprocessed mRNAs would be exported to the cytoplasm and then translated into toxic proteins. Thus, the separation of RNA synthesis in the nucleus and translation in the cytoplasm, and the associated quality control mechanisms that go along with this separation, reduces some of the harmful side-effects of non-functional RNAs that are transcribed from non-functional DNA. This is why both junk DNA and a low level of spurious transcription are tolerated in most eukaryotes (Palazzo and Gregory, 2014; Palazzo and Lee, 2015).

Importantly, non-functional RNAs, whose harmful effects are reduced by eukaryotic quality control systems, are not effectively eliminated by natural selection and some of these can eventually evolve into functional long non-coding RNAs (lncRNAs). These add to the repertoire of bio-active polymers that organisms can use to regulate growth, homeostasis and development. Although some lncRNAs function in the cytoplasm, most operate in the nucleus (Djebali et al., 2012; Tilgner et al., 2012; Derrien et al., 2012; Palazzo and Gregory, 2014; Kaewsapsak et al., 2017). As a result, lncRNAs must be appropriately sorted to allow for their proper retention in the nucleus or export to the cytoplasm. This separation is critical as alterations in mRNA nuclear retention and cytoplasmic export have been associated with various diseases (Borden and Culkovic-Kraljacic, 2018; Bovaird et al., 2018). Furthermore, many neuropathological states are associated with the formation of RNA-protein liquid-liquid phase separated structures that can disrupt proper nuclear-cytoplasmic trafficking by soaking up nuclear transport factors and components of the nuclear pore complex (Mahboubi et al., 2013; Zhang et al., 2018).

So how does this all work? First, long and short RNAs are generally treated very differently. In mammals, RNA length appears to be evaluated by hnRNP C (McCloskey et al., 2012), with transcripts that are shorter

than 200 nucleotides (e.g., snRNAs, tRNAs and miRNAs) being directed towards specialized export pathways (Masuyama et al., 2004; Fuke and Ohno, 2008; McCloskey et al., 2012), while longer RNAs (mRNA and lncRNAs) being shunted to a more generalized pathway that require the major export complex, TREX, and its heterodimeric nuclear transport receptor composed of Nxf1/TAP and Nxt1/p15 (Figure 1) (Katahira et al., 1999; Strässer et al., 2002). In addition to these major export factors, other export-promoting complexes exist. SR proteins, which promote splicing, also help to recruit Nxf1/TAP to these RNAs (Huang et al., 2003; Müller-McNicoll et al., 2016). TREX2, which is thought to localize to the nucleoplasmic side of the nuclear pore, also plays a major role in promoting export (Wickramasinghe et al., 2010; Umlauf et al., 2013; Zhang et al., 2014b; Wickramasinghe et al., 2014). Dbp5, Rae1/Gle1, and Gle2, which associate with the cytoplasmic face of the nuclear pore, may be involved in recycling nuclear export factors back into the nucleoplasm (Blevins et al., 2003; Lund and Guthrie, 2005; Alcázar-Román et al., 2006; Weirich et al., 2006), although this is not quite understood. For reviews on these factors and complexes see (Palazzo and Akef, 2012; Katahira, 2012; Heath et al., 2016; Borden and Culkovic-Kraljacic, 2018). The mechanism that dictates nuclear retention is less well understood. Some of the factors involved are described in later sections of this review.

So how are exported RNAs, which mostly code for protein, differentiated from nuclear retained RNAs, which are typically non-coding? Ultimately these two types of RNAs must differ in one or more ways. This can include *cis*-elements (i.e., particular RNA motifs) or general features such as splicing, polyadenylation, and RNA modifications. These differences will dictate what proteins are loaded onto the RNA, resulting in the formation of a ribonucleoprotein (RNP) complex that determines the ultimate fate of the transcript.

In this review we shall cover what is known about the RNA features that impact the nuclear retention and cytoplasmic export of mRNAs and lncRNAs. However, before we start, there are a few points to keep in mind. First, although we speak of a given RNA species as being retained in the nucleus or exported to the cytoplasm, few RNAs are completely nuclear or cytoplasmic at steady state. Instead, each RNA species exists at some point along a spectrum between these two extremes. Second, the ultimate steady state distribution of an RNA is dictated not only by the rate of RNA export, but also by the rates of RNA synthesis and of RNA decay in both the nucleus and the cytoplasm. Few studies have taken all these various factors into account with some exceptions. Having said that, it is clear that nuclear retention and cytoplasmic export play critical roles in dictating the ultimate distribution of any RNA species. Third, although it is generally true that most mRNAs are well exported, many are not (Djebali et al., 2012; Bahar Halpern et al., 2015; Bouvrette et al., 2018). Likewise, although there is a general consensus that many lncRNAs are nuclear, it

is also clear that several are cytoplasmic, with some studies suggesting that the number of cytoplasmic lncRNAs may be higher than previously thought (Wilk et al., 2016; Bouvrette et al., 2018).

2. What is the default pathway?

Before addressing the question of what sequence determinants impact nuclear export, it becomes necessary to determine whether an RNA which lacks any distinguishing feature is a substrate for nuclear export. In other words, what is the default pathway -- nuclear retention or cytoplasmic export? Three pieces of evidence point to the fact that long RNAs do not require any specialized *cis*-element for them to be exported from the nuclei of mammalian tissue culture cells.

2.1 Reporter mRNAs

Whether the default pathway for any given long RNA was nuclear retention or cytoplasmic export was up for debate for a number of years, due largely to differences between the nuclear/cytoplasmic distribution of mRNAs derived from different reporter genes (Luo and Reed, 1999; Lu and Cullen, 2003; Nott et al., 2004; Masuyama et al., 2004; Palazzo et al., 2007; Valencia et al., 2008; Takemura et al., 2011; Lei et al., 2011, 2013; McCloskey et al., 2012). For example, it had been observed that certain reporter mRNAs transcribed from cDNAs were not exported, suggesting that in the absence of splicing, mRNAs are nuclear retained (Valencia et al., 2008). This confusion was largely due to the fact that it was unclear whether any particular reporter is truly devoid of *cis*-elements or other distinguishing features that may promote or inhibit mRNA nuclear export. More recently, we have demonstrated that two widely used reporters, a mini gene derived from the Drosophila *fushi tarazu* gene (*ftz*), and the *β-globin* mRNA, each have nuclear retention elements (Lee et al., 2015; Akef et al., 2015). Importantly, when the newly identified nuclear retention elements were removed, RNAs generated from these reporters were well exported despite the fact that they are not spliced.

2.2 mRNAs with random sequences

In other experiments it was found that RNAs generated from artificial genes, purported to have “random” sequences, were not exported but were instead rapidly degraded (Dias et al., 2010). One potential problem

with completely random sequences is that they contain elevated numbers of CG dinucleotides, which are depleted in vertebrate genomes (Karlin and Mrázek, 1997). In DNA, CG dinucleotides are often methylated, and when these N^5-methylcytosines undergo spontaneous deamination they are converted to thymidine causing CG dinucleotides to be mutated away in vertebrates (Lindahl, 1993). In contrast, unmethylated cytosines deaminate to uracils, which are efficiently removed by uracil-DNA glycosylase and reconverted back to cytosines. Recently it was found that RNAs with significant numbers of CG dinucleotide are substrates for decay, which would effectively prevent their accumulation in the cytoplasm (Takata et al., 2017). This process likely evolved to protect cells against viral infection. Interestingly, the proteins involved in this decay, ZAP/ZC3HAV1 and TRIM25, are primarily cytoplasmic and are known to be involved in viral RNA degradation.

In the study by Dias and colleagues, the RNA reporters with "random" sequence were in fact generated from the reverse compliment sequences of intronless genes from the human genome (*IFNA1*, *IFNB1* and *HSPB3*). As expected, the three constructs have relatively low CG-content (as is true for almost all human-derived DNA); however, all three are predicted to have either 5' splice site motifs or 3' splice site motifs [scoring ≥ 0.97 according to NNSPLICE 0.9 (Reese et al., 1997)]. These motifs are known to inhibit nuclear export if they are not used for splicing (see sections 5.1-5.2). It is also possible that these transcripts were spliced and that the researchers were detecting the distribution of lariat introns in their experiments. Additionally, it is conceivable that these RNAs may have other nuclear retention elements. Again, interpreting experiments with "random" RNAs is difficult, as unidentified *cis*-elements may drastically alter the behaviors of these transcripts.

2.3 LncRNAs

The last piece of evidence which suggests that nuclear export is the default pathway is that when nuclear localized lncRNAs were analyzed, it was observed that they contained nuclear retention elements (Miyagawa et al., 2012; Zhang et al., 2014a; Lubelsky and Ulitsky, 2018; Shukla et al., 2018). When these nuclear retention elements were removed or mutated, the altered lncRNAs were exported. In one extreme case the intronless *MALAT1* lncRNA was expressed as a series of small fragments (each ~ 1kb), with the majority of the resulting RNAs being efficiently exported (Miyagawa et al., 2012). This allowed researchers to identify two regions that retain this lncRNA in the nucleus by targeting it to nuclear speckles. Moreover, fusion of either of these two nuclear retention fragments to reporters promotes their nuclear retention

(Shukla et al., 2018; Lubelsky and Ulitsky, 2018). Thus, it is likely that lncRNAs like *MALAT1* must be actively retained in the nucleus, and in the absence of these factors, the resulting RNAs are automatically exported to the cytoplasm.

Taking in all of these lines of evidence, it is likely that in the absence of any active *cis*-element, a stable RNA that is capped and polyadenylated is a substrate for nuclear export.

3. The role of RNA processing in nuclear retention and mRNA export

In eukaryotes, most functional RNAs are extensively processed. Although very strong processing signals are found in regions of the genome that are used to produce functional RNA transcripts (be they mRNAs or lncRNAs), weaker processing signals are found throughout the genome. Even comparing mRNAs and lncRNAs, the former are typically more efficiently spliced than the latter (Tilgner et al., 2012; Melé et al., 2017; Mukherjee et al., 2017; Deveson et al., 2018). Thus, robust processing is typically a good indication that the RNA transcript in question is functional and likely encoding a protein (Palazzo and Akef, 2012; Palazzo and Lee, 2015). Moreover, many RNA processing machineries directly interact with, and promote the recruitment of, RNA nuclear export factors. This "coupling" between RNA processing and RNA nuclear export has been extensively documented in other reviews (Maniatis and Reed, 2002; Moore and Proudfoot, 2009; Palazzo and Akef, 2012).

3.1 Splicing

Splicing involves the removal of introns by the spliceosome, which in turn can deposit factors onto the newly spliced RNA. By comparing the localization of these spliced RNAs to transcripts synthesized from cDNAs (which lack introns), it has been observed that splicing in some scenarios enhances the extent and the rate of nuclear export (Luo and Reed, 1999; Palazzo et al., 2007; Valencia et al., 2008). The spliceosome directly interacts with many key mRNA nuclear export factors, such as the TREX component UAP56 (Fleckner et al., 1997; Strässer and Hurt, 2001). Indeed, splicing is known to help recruit TREX components to RNAs (Masuda et al., 2005; Dufu et al., 2010; Chi et al., 2013). This phenomenon is, however, not universal. Most cDNA-derived RNAs (which lack introns) are well exported (Palazzo and Akef, 2012) and can recruit TREX and Nxf1/TAP (Taniguchi and Ohno, 2008; Hautbergue et al., 2009; Akef et al., 2013,

2015; Lee et al., 2015). Likely, where splicing matters most is in transcripts that happen to have nuclear retention elements. In some cases, splicing can override their activity (Akef et al., 2015), while in other cases it cannot (Lee et al., 2015). The second scenario is probably true for lncRNAs that are efficiently spliced and yet still retained in the nucleus (Hacisuleyman et al., 2014).

3.2 5' Capping

The 5' RNA cap is an N^7-methylguanine connected via a 5′ to 5′ triphosphate linkage to the beginning of RNAs which are generated by RNA Polymerase II. This structure recruits the nuclear cap binding complex (CBC), which consists of CBP20 and CBP80. It has been reported that CBP80 can recruit nuclear export factors, such as the TREX component Aly, to the 5'end of spliced (Cheng et al., 2006) and intronless (Nojima et al., 2007) transcripts. More recently it was shown that a functional paralog of CBP80, NCBP3, also interacts with components of the TREX and exon junction complexes (Gebhardt et al., 2015). Importantly, the co-depletion of CBP80 and NCBP3 inhibits mRNA nuclear export (Gebhardt et al., 2015). As such it is clear that the 5' cap is a major contributor to the proper export of mRNAs. Whether it is absolutely required is a bit unclear. The incorporation of non-canonical caps (trimethyl-guanosine [3mGpppG], adenosine [ApppG]) does not block the export of certain microinjected intronless RNAs, but does block the export of intron-containing mRNAs (Palazzo et al., 2007). Since the 5'cap is also required for splicing (Izaurralde et al., 1994), it is possible that RNAs with cap analogues are inefficiently spliced and are thus actively retained in the nucleus. As detailed below, RNA motifs that are associated with introns are potent nuclear retention signals. Lastly, it has been reported that the export of circular RNAs requires UAP56 (Huang et al., 2018), a core factor of the TREX complex that is required for the export of most mRNAs (Strässer and Hurt, 2001; Luo et al., 2001; Kapadia et al., 2006). This would suggest that TREX-mediated export does not strictly require a 5' cap to function.

3.3 3' Cleavage and Polyadenylation

The 3' end of an RNA Polymerase II-generated transcript is recognized and processed by the cleavage and polyadenylation complex (Chan et al., 2011). Members of this complex interact with Aly (Johnson et al., 2009; Shi et al., 2017), the TREX component THOC5 (Katahira et al., 2013; Tran et al., 2014) and Nxf1/TAP (Ruepp et al., 2009). In line with these studies, the RNAs produced from reporter genes with

defective 3' cleavage signals are restricted to the nucleus (Dias et al., 2010). It is however likely that these RNAs are never released from RNA polymerase due to the lack of cleavage, complicating the interpretation of this observation. In another set of experiments, it was found that microinjected RNAs that lack a poly(A)-tail, but are modified at their 3'end to protect the RNAs from degradation are retained in the nucleus (Akef et al., 2013). Again, it is possible that the modification itself, a dialdehyde formed by the oxidation of the free 3' end ribose by periodate, may trigger nuclear retention. On the flip side, a GFP reporter that lacks a tail and contains a 3' terminal triple helix structure derived from the *MALAT1* lncRNA, which stabilizes unpolyadenylated transcripts, is efficiently exported (Wilusz et al., 2012). This observation suggests that either the poly(A)-tail is not strictly required for mRNA export or that this triple helix motif promotes nuclear export, although this element is derived from *MALAT1*, a nuclear lncRNA. Another observation that suggests that the poly(A)-tail is not absolutely required for export is that circular RNAs, which lack a tail, are efficiently exported in a UAP56-dependent manner (Huang et al., 2018). Finally, histone mRNAs, which do not have a poly(A)-tail, are exported by Nxf1 and do not appear to have any export-promoting *cis*-elements (Erkmann et al., 2005).

In summary, it is likely that RNA processing helps to promote export; however, results from a variety of case studies (cDNA derived reporters, GFP mRNA with a 3' terminal triple helix, and circular RNAs) suggest that these processes are not absolutely required. Again, as most RNAs exist on a spectrum between being fully nuclear and being fully cytoplasmic, RNA processing events may help to move the RNA closer to the cytoplasmic end of this continuum.

4. The role of RNA nucleotide modifications in nuclear retention and mRNA export

It has been known for quite some time that RNA is extensively modified; however, until recently the majority of these studies focused on these modifications within tRNA and rRNA. More recently it has been observed that mRNA and lncRNAs are also modified. Furthermore, some of these modifications appear to impact nuclear export.

4.1 Adenosine to Inosine Editing

Adenosine to inosine editing was the first RNA modification known to affect nuclear export. Specifically, it was observed that double stranded RNA (dsRNA) was a substrate for the RNA specific adenosine deaminase (ADAR), which converts adenosine to inosine (Polson et al., 1991). This reaction occurs specifically in the nucleus and promotes the nuclear retention of these RNAs (Zhang and Carmichael, 2001). Thus RNAs that are prone to forming long dsRNA, including mRNAs with inverted Alu repeats and viruses (Kumar and Carmichael, 1997; Levanon et al., 2004; Athanasiadis et al., 2004; Kim et al., 2004; Blow et al., 2004), are modified and retained in paraspeckles (Chen et al., 2008). In certain cases nuclear retention of inosine-containing mRNAs can also be used to regulate gene expression (Prasanth et al., 2005). Interestingly, this nuclear retention pathway appears to be less active in human embryonic stem cells due to the fact that they do not express the lncRNA *NEAT1*, which is required for paraspeckle formation (Chen and Carmichael, 2009).

4.2 Other RNA modifications

In the last six years, it has become clear that other modifications, which were known to occur in tRNA and rRNA, play significant roles in mRNA biology. This includes N^6-methyladenosine (Dominissini et al., 2012; Meyer et al., 2012), which accumulates near the stop codon, N^5-methylcytosine (Squires et al., 2012) and N^1-methyladenosine (Dominissini et al., 2016; Li et al., 2016a), which both accumulate near the start codon, and pseudouridine (Carlile et al., 2014; Schwartz et al., 2014; Li et al., 2015), which accumulates in the ORF and 3'UTR.

Recently, it has been reported that N^6-methyladenosine promotes the nuclear export of mRNAs (Roundtree et al., 2017) through the action of the YTHDC1 protein, which directly binds to the modified base and helps to recruit nuclear export factors to the mRNA. This makes sense as depletion of the N^6-methyladenosine demethylase, ALKBH5, enhances overall mRNA export (Zheng et al., 2013). Similarly, N^5-methylcytosine has also been reported to promote mRNA nuclear export by recruiting Aly to the transcript (Yang et al., 2017). Although N^1-methyladenosine has not been directly linked to export, this modification is enriched in the 5' terminal exon of a particular class of mRNAs (Cenik et al., 2017). These mRNAs have interesting 5' terminal exons. Not only are they modified, but they also tend to contain the start codon (in most human genes the start codon is found in internal exons), and are enriched in certain GC-rich motifs that are associated with exon junction complexes (Singh et al., 2012). Typically, exon junction complexes are deposited upstream of all newly formed exon-exon splice sites; however, in a subset of genes the exon

junction complex also associates with these GC-rich motifs. Importantly, this complex has also been found to bind to nuclear export factors (Le Hir et al., 2001; Singh et al., 2012), although it is not strictly required for export (Palazzo et al., 2007).

In conclusion, RNA modifications that have been reported to promote export may enhance this process, especially if an RNA has nuclear retention elements; however, it is likely that RNA modifications are not absolutely required to promote export.

5. The role of cis-elements in nuclear retention and mRNA export

5.1 The 5' Splice Site Motif

Some of the most studied RNA motifs that affect the distribution and stability of mature mRNA are the 5' and 3' splice site motifs, which specify the boundaries of introns. These are typically removed by the act of splicing. Importantly, these motifs are found in fully processed exported RNAs of many viruses, such as HIV. In its normal life cycle, HIV produces both spliced and unspliced RNAs from the same primary transcript, the latter being used to make late-stage proteins and to generate the RNA-based genome that will be incorporated into new viruses that are assembled in the cytoplasm of the host cell. Importantly, these unspliced RNAs are retained in the nucleus in early stages by the presence of intronic sequences (Chang and Sharp, 1989; Lu et al., 1990; Borg et al., 1997; Séguin et al., 1998). These retention signals can be overcome in late stages by the virally encoded Rev protein, which recognizes the Rev response element, an RNA structure that is present in the late stage RNAs and the viral RNA genome (Emerman et al., 1989; Chang and Sharp, 1989; Tan et al., 1996). In the absence of Rev, the nuclear retention of these RNAs was mediated in part by U1 snRNP, the component of the spliceosome that recognizes the 5' splice site motif (Lu et al., 1990) (Figure 2). It should be noted that the Rev response element itself also contributes to the nuclear retention of the late-stage viral mRNAs and of the HIV genomic RNA when Rev protein is not present (Brighty and Rosenberg, 1994; Nasioulas et al., 1994).

In other work, it was also demonstrated that when the 5' splice site motif was present in the terminal exon of an mRNA, it inhibited expression of the encoded protein. This was due in part to the fact that this element suppresses 3' polyadenylation, which in turn targets the mRNA for degradation (Gunderson et al., 1998)

(Figure 2). This configuration is not only seen in certain viral mRNAs, but also in human mRNAs. For example, a mutation in the LAMTOR2 gene, which is associated with congenital neutropenia, creates a novel 5' splice site in the 3' UTR that results in the inhibition of gene expression (Langemeier et al., 2012). Importantly, this inhibition is likely due to the recruitment of U1 snRNP to the mature mRNA (Langemeier et al., 2012), through the direct hybridization of the U1 snRNA with the 5' splice site. Indeed, when the sequence of the U1 snRNA is altered so that it now base pairs to some other mRNA, these newly targeted transcripts becomes silenced (Fortes et al., 2003; Abad et al., 2008; Goraczniak et al., 2009; Blázquez and Fortes, 2013). A protein component of the U1 snRNP, U1-70K, is required for this inhibition by directly interacting and inhibiting poly(A)-polymerase (Gunderson et al., 1998).

As we stated in the introduction, disentangling the effects of mRNA stability and nuclear retention on the final nuclear/cytoplasmic distribution of an mRNA can be challenging. This is certainly the case with the 5' splice site motif which appears to promote both RNA degradation and RNA nuclear retention. To tease these two forces apart, we monitored the level and distribution of newly synthesized reporter mRNAs that contained or lacked a 5' splice site motif in its 3'UTR. This was accomplished by microinjecting DNAs that were transcribed into each mRNA species, then allowing transcription to proceed for a short amount of time (15 to 20min) before halting transcription with α-amanitin, and then monitoring the newly transcribed RNA by fluorescence in situ hybridization at various timepoints after injection. Using this approach we found that about half of newly synthesized reporter mRNAs that contained a 5' splice site motif are rapidly degraded, with the remaining fraction being retained in the nucleus as polyadenylated RNAs (Lee et al., 2015) (Figure 2). Interestingly, the nuclear retained RNAs accumulate in nuclear speckles, subnuclear regions where post-transcriptional splicing is thought to occur (Dias et al., 2010; Vargas et al., 2011). Indeed, unspliced mRNAs which are generated by the inhibition of the U2 or U4 snRNPs, also accumulate in nuclear speckles (Kaida et al., 2007; Hett and West, 2014). These unspliced RNAs and 5' splice site bearing RNAs are likely targeted to nuclear speckles by U1. Then, the subsequent failure to complete the splicing reaction prevents these RNAs from exiting the nuclear speckles. In agreement with these results, the artificial tethering of U1-70K to a reporter RNA prevents its nuclear export, although the authors did not test for nuclear speckle targeting (Takemura et al., 2011). Surprisingly, 5' splice site motif-containing mRNAs are still able to recruit UAP56 and Nxf1/TAP (Lee et al., 2015), suggesting that if they could reach the pore, these mRNAs could cross it; however, access to the pore may be prevented by their sequestration into speckles (Figure 2). This may explain why many poorly exported mRNAs are also localized to nuclear speckles (Bahar Halpern et al., 2015).

The presence of 5' splice site motifs may also be critical for the nuclear retention of many lncRNAs and may help to distinguish them from mRNAs. According to annotated databases of human genes, fully mature lncRNAs, unlike mature mRNAs, are not depleted of 5' splice site motifs in their terminal exons (Lee et al., 2015). Even when comparing intronless RNAs, lncRNAs have higher levels of 5' splice site motifs than mRNAs (Lee et al., 2015). These numbers may be an underestimate as lncRNAs are not as efficiently spliced as mRNAs (Tilgner et al., 2012; Melé et al., 2017; Mukherjee et al., 2017; Deveson et al., 2018), with many isoforms containing retained introns due to the inefficient recruitment of spliceosomal factors to 3' splice sites (Melé et al., 2017). Interestingly, the corresponding 5' splice sites of these inefficiently spliced introns still recruit U1 (Melé et al., 2017), and thus likely promote nuclear retention. Indeed, lncRNA splicing appears to be sloppier than mRNA splicing, with each lncRNA gene producing a multitude of different isoforms with altered splice junctions (Deveson et al., 2018), and this may also cause the inclusion of 5' splice site motifs into the mature RNA.

The 5' splice site motif also inhibits 3' cleavage. When cells were depleted of U1 snRNPs, prematurely truncated mRNAs started to appear (Kaida et al., 2010; Berg et al., 2012; Almada et al., 2013). This was due to a decrease of splicing which led to the appearance of retained introns, which in turn contained cryptic 3' cleavage/polyadenylation sites that were inappropriately used by the 3' cleavage machinery. Importantly, these truncated RNAs contain intact 5' splice site upstream of the new 3' end. It was inferred that under normal circumstances the binding of U1 inhibits 3' cleavage from any sites in the downstream intron (Figure 3). This finding is in agreement with studies of Bovine Papilloma Virus and HIV RNAs where the recruitment of U1 to a 5' splice site inhibited proximal 3' cleavage events (Furth et al., 1994; Ashe et al., 1995, 1997; Vagner et al., 2000). Similar results were seen with the mutant form of the LAMTOR2 mRNA (Langemeier et al., 2012). It should be noted that in normal situations, suppression of 3' cleavage by U1 snRNP helps to perform two tasks: first it represses the misprocessing of mRNAs by preventing the activity of cryptic 3' cleavage/polyadenylation sites that are found in introns; secondly, it enforces promoter directionality. In particular, it was found that in bidirectional promoters which generate an unstable short cryptic transcript in one direction and a stable protein-coding mRNA in the other direction, that 3' cleavage/polyadenylation consensus sites were enriched in the former, and 5' splice site motifs were enriched in the latter (Almada et al., 2013) (Figure 3). Under normal conditions the transcriptional elongation of these cryptic transcripts is curtailed by the presence of these 3' cleavage/polyadenylation sites. Early 3' cleavage promotes RNA degradation, although the exact mechanism in vertebrates remains unclear (Proudfoot, 2016). In contrast, the 5' splice site motif present on the opposite transcriptional unit prevents the utilization of any downstream cryptic 3'cleavage/polyadenylation site and thus promotes the

transcriptional extension and stability of functional RNAs. This arrangement of 5' splice site motifs and 3' cleavage/polyadenylation sites is sometimes referred to as the U1-PAS axis (PAS stands for polyadenylation sites).

One important complex which may promote nuclear retention and degradation of RNAs that contain 5' splice site motifs is the PAXT complex, which consists of the RNA helicase, Mtr4, the zinc finger-containing protein, ZFC3H1, and the nuclear poly(A) binding protein, PABPN1 (Meola et al., 2016). Depletion of Mtr4 or ZFC3H1 resulted in the cytoplasmic accumulation of truncated mRNAs that utilized cryptic 3' cleavage sites from intronic regions (Ogami et al., 2017). Mtr4 may promote nuclear retention of these transcripts by competing with the RNA export adaptor Aly for binding of the 5' cap (Fan et al., 2017). Furthermore, Mtr4 is also a co-activator of the nuclear exosome (Schilders et al., 2007), the major RNA degradation machinery in the nucleus, suggesting that PAXT may also target these RNAs for degradation. It is currently unclear how the PAXT complex would recognize its substrates, although one possibility is that it interacts with U1 that is bound to misprocessed mRNAs.

In budding yeast, mRNAs with unspliced introns are also nuclear retained and degraded, and this likely requires an intact 5' splice site (Legrain and Rosbash, 1989). This retention activity requires the Mlp1/2 proteins (Galy et al., 2004; Vinciguerra et al., 2005), which form the nuclear basket, a structure that sits on the nucleoplasmic face of the nuclear pore complex. In vertebrates, the nuclear basket protein TPR, which shares some homology with Mlp1/2, is also required for the nuclear retention of intron-bearing mRNAs (Coyle et al., 2011; Rajanala and Nandicoori, 2012). Interestingly, TPR is also required for mRNA export (Shibata et al., 2002; Umlauf et al., 2013; Wickramasinghe et al., 2014), likely by associating with the TREX2 factor GANP (Figure 1), which is essential for the nuclear export of most mRNAs (Wickramasinghe et al., 2010; Zhang et al., 2014b).

Finally, it should be pointed out that the nuclear retention of mRNAs harboring retained introns may also be used to regulate gene expression. It has been found that certain regulated mRNAs contain "detained" introns that are poorly spliced, leading to the retention of the transcripts into the nucleus (Boutz et al., 2015; Mauger et al., 2016; Naro et al., 2017). These are typically the last introns in the pre-mRNA, and it is likely that the primary signal for nuclear retention is the presence of a 5' splice site motif in these terminal exons. These retained mRNAs are stable and not subject to degradation. However, in response to some signal,

these introns are post-transcriptionally spliced, releasing the mRNAs from the nucleus and triggering protein production.

5.2 Other intron-associated motifs

Besides the 5' splice site motif, it has been reported that other sequences that are normally associated with introns also potentiate nuclear retention. Typically, the 3' end of an intron is defined by a polypyrimidine track which can be recognized by the polypyrimidine track binding protein (PTB). Association of PTB with mature RNAs is known to inhibit splicing and nuclear export (Yap et al., 2012; Roy et al., 2013). In addition, the 3' end of the intron also recruits the splicing factor U2AF65, whose association with a mature RNA also promotes nuclear retention (Takemura et al., 2011). Finally, it also appears that the presence of an intact branch-point sequence in the mature mRNA also promotes nuclear retention in budding yeast (Legrain and Rosbash, 1989; Rain and Legrain, 1997). Thus, it is likely that several different intron-associated elements may help to promote the nuclear retention and decay of RNA.

5.3 Transposable element associated motifs

The majority of the human genome is composed of dead transposable elements, constituting half to two-thirds of all DNA (Gregory, 2005; de Koning et al., 2011). Although they are numerous, they are rarely found in mature mRNAs and found at moderate levels in lncRNAs (Kelley and Rinn, 2012). When they are present, they usually inhibit nuclear export and promote RNA decay. As described above, if a pair of transposable elements are found in the sense and anti-sense orientation in a single transcript, they can hybridize to form double stranded RNAs. These regions either become substrates for the ADAR enzyme and thus acquire inosine modifications (Chen et al., 2008; Chen and Carmichael, 2008), or are recognized by the RNA binding protein Staufen, which targets these RNAs for decay (Gong and Maquat, 2011; Park and Maquat, 2013; Elbarbary et al., 2013; Lucas et al., 2018). In addition, double stranded RNAs activate the kinase PKR, which then phosphorylates the translation initiation factor eIF2α and thus shuts down global translation (Clemens and Elia, 1997). Typically, PKR is activated by double stranded viruses, however it is also known to regulate the processing of certain host mRNAs (Ilan et al., 2017). It remains unclear if PKR activity impacts nuclear export.

It is likely that other features associated with transposable elements are recognized by nuclear retention machinery. It was recently found that the reverse complement of the Alu SINE, a primate-specific transposable element, contains a 42 nucleotide long element, named SIRLOIN, that mediates nuclear retention by recruiting the RNA binding protein, hnRNP K (Lubelsky and Ulitsky, 2018). A similar C-rich motif that contributed to nuclear retention was found in a large analysis of human lncRNAs (Shukla et al., 2018). Since Alu elements are not found outside of primates, lncRNAs must use other elements, especially in non-primates. In addition, it appears that many transposable elements are recognized by particular C2H2 zinc finger proteins (Emerson and Thomas, 2009; Rowe and Trono, 2011; Schmitges et al., 2016), many of which contain not only the capability to bind DNA, but also RNA (Burdach et al., 2012). It has been speculated that when a new transposable element invades a genome, it catalyzes the evolution of novel zinc finger proteins that protects the host. These zinc finger proteins likely repress transposable element activity primarily through transcriptional silencing, although it is also possible that these proteins may help target RNAs for decay or nuclear retention.

5.4 Other cis-elements that promote nuclear retention

A few other *cis*-elements that promote nuclear retention have been characterized in the literature. As mentioned above, the Rev responsive element promotes nuclear retention (Brighty and Rosenberg, 1994; Nasioulas et al., 1994). Another example is the AGCCC motif which promotes the nuclear retention of the *BORG* lncRNA (Zhang et al., 2014a). Although the authors of this study show that the presence of this motif correlated with the nuclear/cytoplasmic distribution of a few lncRNAs and mRNAs, such a sequence would be predicted to be depleted from mRNAs in general; however, in a large genome-wide analysis, we have failed to detect such a depletion (A. F. Palazzo, unpublished observations). This is unlike the 5' splice site motif, which is depleted from intronless mRNAs and the 3' terminal exons of human mRNAs (Lee et al., 2015).

In many cases, recruitment of certain proteins to the RNA has been linked to nuclear retention, however it remains unclear whether the simple presence of their RNA-binding motifs promotes retention more broadly throughout the transcriptome. This is true of the *Firre* lncRNA, whose nuclear retention requires the recruitment of hnRNP U protein (Hacisuleyman et al., 2014). Similarly, it has been reported that the recruitment of hnRNP A2 inhibits nuclear export (Lévesque et al., 2006). Again, a more global analysis of

how these factors affect the nuclear/cytoplasmic distribution of all RNAs would be useful in determining whether other nuclear retention elements exist.

Other global analyses have tried to identify nuclear retention/export motifs by sequencing RNA derived from nuclear and cytoplasmic compartments (Bahar Halpern et al., 2015; Bouvrette et al., 2018). Interestingly, both studies found a reasonable number of mRNAs that were poorly exported. Although the distribution of mRNAs with either the nuclear or cytosolic compartment corelated with the association of certain RNA binding proteins, no obvious patterns were discovered. This is in contrast to lncRNAs where the presence of motifs that are either associated with transposable elements (Lubelsky and Ulitsky, 2018; Shukla et al., 2018) or unused splicing signals (Lee et al., 2015; Melé et al., 2017) likely promote widespread nuclear retention. Why would lncRNAs and mRNAs have different mechanisms for their nuclear distribution? One difference may be that nuclear lncRNAs are actively retained while nuclear mRNAs are simply exported to the cytoplasm at a very low rate. This would allow these particular mRNAs to accumulate in the nucleus at high levels. It has been hypothesized that since these large pool of nuclear mRNAs would slowly exit the nucleus, they would supply the cytoplasm with a steady level of mRNA over long periods of time and this could help to buffer the protein translation machinery in the cytoplasm from any wide fluctuations in mRNA production in the nucleus (Bahar Halpern et al., 2015). This may be especially important for genes that experience transcriptional bursts, the sporadic production of many mRNAs in a short interval, followed by periods of inactivity (Larson, 2011). Without this buffering, mRNA levels in the cytoplasm would stochastically increase and decrease over short intervals of time, especially if the mRNA has a short half-life.

A few studies have uncovered large RNA elements that have nuclear retention activity but remain ill-defined. Two of the best examples are the intronless *β-globin* mRNA and the *MALAT1* lncRNA. In the case of *β-globin*, the nuclear retention activity maps to the last 210 nucleotides of the open reading frame (Akef et al., 2015). This nuclear retention activity can be overcome by either extending the length of the transcript (Akef et al., 2015), including an intron to promote splicing (Valencia et al., 2008; Akef et al., 2015), or by inserting certain export-promoting viral RNA elements (Guang and Mertz, 2005; Chi et al., 2014). Deleting the first or the second half of this 210 nucleotide region does not disrupt nuclear retention, suggesting that there may be multiple sequences that account for this activity. Despite this, the two halves do not share any obvious motif or structure. In the case of *MALAT1*, its two nuclear speckle targeting regions (termed regions "E" and "M") are also ill-defined (Miyagawa et al., 2012). In the case of region E which is about 1KB in length, elimination of the first or last half disrupts its activity. For region M, its activity maps to 600

nucleotides, but it is disrupted if it is truncated any further. It is likely that the *XIST*, *NEAT1* and *TUG1* lncRNAs also have large nuclear retention elements (Lubelsky and Ulitsky, 2018; Shukla et al., 2018). Ultimately, it remains possible that these pieces of RNA contain one or more discrete motifs or structures that have weak nuclear retention activity (Shukla et al., 2018), and that further in-depth studies would be needed in order to better define these elements.

5.5 Cis-elements that promote nuclear export

Many viral elements are known to promote nuclear export; however, a number of these act to overcome nuclear retention elements such as the presence of unspliced introns. Besides the Rev responsive element (described above), the most well studied is the constitutive transport element (CTE) of type D retroviruses (Bray et al., 1994). This large-structured RNA directly recruits Nxf1/TAP to the transcript (Grüter et al., 1998). Interestingly, the *Nxf1* mRNA contains a CTE-like element that can also recruit Nxf1/TAP (Li et al., 2006; Wang et al., 2015). These elements appear to modulate the export of *Nxf1* mRNA isoforms that contain a retained intron (and hence a 5' splice site). The mRNA is then translated into a short isoform of Nxf1 that may play a role in mRNA trafficking (Li et al., 2016b).

Some mRNAs have been described to have *cis*-elements that promote nuclear export. mRNAs that encode proteins required for the cell cycle, contain an export promoting element in their 3'UTR which consists of a stem loop structure that recruits the eIF4E protein (Culjkovic et al., 2005, 2006). Intriguingly, the export of these transcripts requires UAP56, but not Nxf1/TAP (Topisirovic et al., 2009). Instead they use the CRM1 nuclear transport receptor, which promotes the export of proteins. The recruitment of HuR to mRNAs and lncRNAs has also been reported to promote their nuclear export (Fan and Steitz, 1998; Noh et al., 2016). Finally, it has been reported that naturally intronless transcripts contain specialized cytoplasmic accumulation region elements (CAR-E), which recruit specific complexes to the RNA (Lei et al., 2011, 2013). Some of the interpretations of these experiments are complicated by the fact that CAR-Es were fused to reporters harboring nuclear retention elements whose activity can be overcome by simply extending the length of the transcript (see discussion in Akef et al., 2015). Notably, the export of these mRNAs require the TREX component UAP56, which appears to be recruited to reporter RNAs that do not contain any known nuclear export elements (Taniguchi and Ohno, 2008; Lee et al., 2015; Akef et al., 2015). Thus, the functional relevance of these purported export-promoting elements seems unclear at this time. It is likely

that *bone fide* export-promoting elements, such as the CTE, function by overcoming the activity of nuclear retention elements, such as the ones present in mRNAs with retained introns.

6. Nuclear retention and export of RNAs, a force for Constructive Neutral Evolution?

6.1 The conversion of junk RNA to lncRNA

The nuclear/cytoplasmic distribution of RNAs plays an important role in shaping the genomic content of eukaryotes by evolution. In particular, the nuclear retention and degradation of spurious transcripts eliminates much of the harm caused by junk RNA and hence reduces the deleteriousness of cryptic transcriptional start sites and intergenic DNA regions that harbor such sites (Palazzo and Akef, 2012; Palazzo and Gregory, 2014; Palazzo and Lee, 2015). As a result, junk DNA and its associated junk RNA are not effectively eliminated by natural selection. It is likely that these non-functional transcripts act as the raw substrates for natural selection and some are converted into novel functional lncRNAs. Thus, in a sense, junk RNA and functional lncRNAs come as a package. The idea that neutral mutations (i.e., intergenic insertions, and the serendipitous creation of cryptic transcriptional start sites) create novel entities (i.e., junk RNA) that are subsequently shaped by natural selection to create novel genes (i.e., lncRNAs) is an example of a general process called *constructive neutral evolution* (Stoltzfus, 1999, 2012; Gray et al., 2010; Lukeš et al., 2011). A key component in this process is the role of the nuclear/cytoplasmic division (and its associated quality control mechanisms) in reducing the deleteriousness of spurious transcription.

So how exactly would junk RNA be converted to lncRNA? Likely, this is a step by step process where new entities are created by non-adaptive processes and then acquire functions which can be selected for by natural selection. One example is presented in Figure 4. First, random mutations create and destroy cryptic transcriptional start sites (TSSs). These sites are engaged by RNA polymerase II which not only generates unstable ncRNAs, but also recruits histone modification enzymes that alter chromatin packaging downstream from the TSS (van Werven et al., 2012; Castelnuovo et al., 2013; Kim et al., 2016; Woo et al., 2017). If the resulting altered histone modifications impart some benefit by regulating nearby genes in a way that improves the fitness of the organism, then the transcriptional event and its cryptic transcriptional start site will be selectively retained. Eventually the ncRNA generated from these loci, which is initially a by-product, may act as a platform to help assemble chromatin remodeling complexes in the vicinity of their

target genes. In this way, the ncRNA acquires a novel function over time and is thus converted into a lncRNA (Figure 4).

This conversion process may frequently occur in tissues that have a high amount of spurious transcription, such as in developing spermatids (Kaessmann, 2010; Jandura and Krause, 2017). During sperm development, DNA is unpackaged from histones and then repackaged into protamines. This transiently exposed DNA can act as a non-specific substrate for RNA polymerases causing high levels of spurious transcription. Once a ncRNAs acquires some associated function in the testes, it can subsequently be expressed in other tissues. This is known as the "out of the testes" hypothesis (Kaessmann, 2010; Jandura and Krause, 2017).

6.2 The conversion of misprocessing to alternative processing

The nuclear/cytoplasmic distribution and degradation of RNA also facilitates the evolution of alternative splicing. In particular, by retaining and degrading misprocessed mRNAs, they are not efficiently translated into proteins and do not cause much harm to the organism. This reduces the deleteriousness of splicing and polyadenylation errors and prevents their elimination by natural selection. This may explain why splicing appears to be inherently sloppy in mammalian cells. In support of this idea, it has been widely noted that although most genes are alternatively spliced, they typically give rise to only one polypeptide (Tress et al., 2017), suggesting that many spliced isoforms are not translated due to their degradation and/or nuclear retention. As such, nuclear/cytoplasmic distribution and degradation of RNA prevents the elimination of cryptic splice site motifs and any other splicing-regulating elements that may appear by random mutation in the genome. These elements then act as the raw substrates necessary for the evolution of functional alternative splicing events. This is another example of constructive neutral evolution in action. In this case the newly created entities are splice sites and/or elements that regulate splicing, which are rendered effectively neutral by the RNA nuclear retention and degradation machinery, and these provide the raw substrates for the evolution of alternatively spliced isoforms. A similar process can be invoked for the evolution of 3' cleavage/polyadenylation sites.

7. Conclusion

Results from ENCODE point to a wide diversity of nuclear/cytoplasmic distribution for many different types of RNA molecules (Djebali et al., 2012; Palazzo and Gregory, 2014). Over the past few years, we have gained a fuller picture of the rules that dictate RNA distribution to these two compartments. We have established that any stable long RNA is a substrate for nuclear export unless it contains a nuclear retention element. Undoubtably, splicing and other RNA processing events further enhance nuclear export. In addition, RNA modifications also play an important role in this process. Although our understanding of the major components that drive export are well known, we still must identify nuclear retention complexes and determine their mode of action to obtain a full picture of how the nuclear and cytoplasmic transcriptomes are achieved.

Acknowledgments

We would like to thank J. Wan and C. Smibert for comments on the manuscript. This work was supported by a grant from the Canadian Institutes of Health Research to A.F.P. (FRN 102725).

Author Contributions Statement

The manuscript was written by A.F.P. and E.S.L.

References:

Abad, X., M. Vera, S.P. Jung, E. Oswald, I. Romero, V. Amin, P. Fortes, and S.I. Gunderson. 2008. Requirements for gene silencing mediated by U1 snRNA binding to a target sequence. *Nucleic Acids Res.* 36:2338–2352. doi:10.1093/nar/gkn068.

Akef, A., E.S. Lee, and A.F. Palazzo. 2015. Splicing promotes the nuclear export of β-globin mRNA by overcoming nuclear retention elements. *RNA*. 21:1908–1920. doi:10.1261/rna.051987.115.

Akef, A., H. Zhang, S. Masuda, and A.F. Palazzo. 2013. Trafficking of mRNAs containing ALREX-promoting elements through nuclear speckles. *Nucleus*. 4:326–340.

Alcázar-Román, A.R., E.J. Tran, S. Guo, and S.R. Wente. 2006. Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export. *Nat. Cell Biol.* 8:711–716. doi:10.1038/ncb1427.

Almada, A.E., X. Wu, A.J. Kriz, C.B. Burge, and P.A. Sharp. 2013. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature*. 499:360–363. doi:10.1038/nature12349.

Ashe, M.P., P. Griffin, W. James, and N.J. Proudfoot. 1995. Poly(A) site selection in the HIV-1 provirus: inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. *Genes Dev.* 9:3008–3025.

Ashe, M.P., L.H. Pearson, and N.J. Proudfoot. 1997. The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. *EMBO J.* 16:5752–5763. doi:10.1093/emboj/16.18.5752.

Athanasiadis, A., A. Rich, and S. Maas. 2004. Widespread A-to-I RNA Editing of Alu-Containing mRNAs in the Human Transcriptome. *PLOS Biol.* 2:e391. doi:10.1371/journal.pbio.0020391.

Bahar Halpern, K., I. Caspi, D. Lemze, M. Levy, S. Landen, E. Elinav, I. Ulitsky, and S. Itzkovitz. 2015. Nuclear Retention of mRNA in Mammalian Tissues. *Cell Rep.* 13:2653–2662. doi:10.1016/j.celrep.2015.11.036.

Berg, M.G., L.N. Singh, I. Younis, Q. Liu, A.M. Pinto, D. Kaida, Z. Zhang, S. Cho, S. Sherrill-Mix, L. Wan, and G. Dreyfuss. 2012. U1 snRNP Determines mRNA Length and Regulates Isoform Expression. *Cell*. 150:53–64. doi:10.1016/j.cell.2012.05.029.

Blázquez, L., and P. Fortes. 2013. U1 snRNP control of 3'-end processing and the therapeutic application of U1 inhibition combined with RNA interference. *Curr. Mol. Med.* 13:1203–1216.

Blevins, M.B., A.M. Smith, E.M. Phillips, and M.A. Powers. 2003. Complex formation among the RNA export proteins Nup98, Rae1/Gle2, and TAP. *J. Biol. Chem.* 278:20979–20988. doi:10.1074/jbc.M302061200.

Blow, M., P.A. Futreal, R. Wooster, and M.R. Stratton. 2004. A survey of RNA editing in human brain. *Genome Res.* 14:2379–2387. doi:10.1101/gr.2951204.

Borden, K., and B. Culkovic-Kraljacic. 2018. mRNA Export and Its Dysregulation in Disease. *In* Nuclear-Cytoplasmic Transport. W. Yang, editor. Springer International Publishing, Cham. 179–204.

Borg, K.T., J.P. Favaro, and S.J. Arrigo. 1997. Involvement of human immunodeficiency virus type-1 splice sites in the cytoplasmic accumulation of viral RNA. *Virology*. 236:95–103. doi:10.1006/viro.1997.8726.

Boutz, P.L., A. Bhutkar, and P.A. Sharp. 2015. Detained introns are a novel, widespread class of post-transcriptionally spliced introns. *Genes Dev.* 29:63–80. doi:10.1101/gad.247361.114.

Bouvrette, L.P.B., N.A.L. Cody, J. Bergalet, F.A. Lefebvre, C. Diot, X. Wang, M. Blanchette, and E. Lécuyer. 2018. CeFra-seq reveals broad asymmetric mRNA and noncoding RNA distribution profiles in Drosophila and human cells. *RNA*. 24:98–113. doi:10.1261/rna.063172.117.

Bovaird, S., D. Patel, J.-C.A. Padilla, and E. Lécuyer. 2018. Biological functions, regulatory mechanisms and disease relevance of RNA localization pathways. *FEBS Lett.* in press. doi:10.1002/1873-3468.13228.

Bray, M., S. Prasad, J.W. Dubay, E. Hunter, K.T. Jeang, D. Rekosh, and M.L. Hammarskjöld. 1994. A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency

virus type 1 expression and replication Rev-independent. *Proc. Natl. Acad. Sci. U. S. A.* 91:1256–1260.

Brighty, D.W., and M. Rosenberg. 1994. A cis-acting repressive sequence that overlaps the Rev-responsive element of human immunodeficiency virus type 1 regulates nuclear retention of env mRNAs independently of known splice signals. *Proc. Natl. Acad. Sci.* 91:8314–8318. doi:10.1073/pnas.91.18.8314.

Burdach, J., M.R. O'Connell, J.P. Mackay, and M. Crossley. 2012. Two-timing zinc finger transcription factors liaising with RNA. *Trends Biochem. Sci.* 37:199–205. doi:10.1016/j.tibs.2012.02.001.

Carlile, T.M., M.F. Rojas-Duran, B. Zinshteyn, H. Shin, K.M. Bartoli, and W.V. Gilbert. 2014. Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature*. 515:143–146. doi:10.1038/nature13802.

Castelnuovo, M., S. Rahman, E. Guffanti, V. Infantino, F. Stutz, and D. Zenklusen. 2013. Bimodal expression of PHO84 is modulated by early termination of antisense transcription. *Nat. Struct. Mol. Biol.* 20:851–858. doi:10.1038/nsmb.2598.

Cenik, C., H.N. Chua, G. Singh, A. Akef, M.P. Snyder, A.F. Palazzo, M.J. Moore, and F.P. Roth. 2017. A common class of transcripts with 5'-intron depletion, distinct early coding sequence features, and N(1)-methyladenosine modification. *RNA*. 23:270–283. doi:10.1261/rna.059105.116.

Chan, S., E.-A. Choi, and Y. Shi. 2011. Pre-mRNA 3'-end processing complex assembly and function. *Wiley Interdiscip. Rev. RNA*. 2:321–335. doi:10.1002/wrna.54.

Chang, D.D., and P.A. Sharp. 1989. Regulation by HIV Rev depends upon recognition of splice sites. *Cell*. 59:789–795.

Chen, L.-L., and G.G. Carmichael. 2008. Gene regulation by SINES and inosines: biological consequences of A-to-I editing of Alu element inverted repeats. *Cell Cycle Georget. Tex*. 7:3294–3301.

Chen, L.-L., and G.G. Carmichael. 2009. Altered Nuclear Retention of mRNAs Containing Inverted Repeats in Human Embryonic Stem Cells: Functional Role of a Nuclear Noncoding RNA. *Mol. Cell*. 35:467–478. doi:10.1016/j.molcel.2009.06.027.

Chen, L.-L., J.N. DeCerbo, and G.G. Carmichael. 2008. Alu element-mediated gene silencing. *EMBO J.* 27:1694–1705. doi:10.1038/emboj.2008.94.

Cheng, H., K. Dufu, C.-S. Lee, J.L. Hsu, A. Dias, and R. Reed. 2006. Human mRNA export machinery recruited to the 5' end of mRNA. *Cell*. 127:1389–1400. doi:10.1016/j.cell.2006.10.044.

Chi, B., K. Wang, Y. Du, B. Gui, X. Chang, L. Wang, J. Fan, S. Chen, X. Wu, G. Li, and H. Cheng. 2014. A Sub-Element in PRE enhances nuclear export of intronless mRNAs by recruiting the TREX complex via ZC3H18. *Nucleic Acids Res.* 42:7305–7318. doi:10.1093/nar/gku350.

Chi, B., Q. Wang, G. Wu, M. Tan, L. Wang, M. Shi, X. Chang, and H. Cheng. 2013. Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic Acids Res.* 41:1294–1306. doi:10.1093/nar/gks1188.

Clemens, M.J., and A. Elia. 1997. The Double-Stranded RNA-Dependent Protein Kinase PKR: Structure and Function. *J. Interferon Cytokine Res.* 17:503–524. doi:10.1089/jir.1997.17.503.

Coyle, J.H., Y.-C. Bor, D. Rekosh, and M.-L. Hammarskjold. 2011. The Tpr protein regulates export of mRNAs with retained introns that traffic through the Nxf1 pathway. *RNA*. 17:1344–1356. doi:10.1261/rna.2616111.

Culjkovic, B., I. Topisirovic, L. Skrabanek, M. Ruiz-Gutierrez, and K.L.B. Borden. 2005. eIF4E promotes nuclear export of cyclin D1 mRNAs via an element in the 3′UTR. *J. Cell Biol.* 169:245–256. doi:10.1083/jcb.200501019.

Culjkovic, B., I. Topisirovic, L. Skrabanek, M. Ruiz-Gutierrez, and K.L.B. Borden. 2006. eIF4E is a central node of an RNA regulon that governs cellular proliferation. *J. Cell Biol.* 175:415–426. doi:10.1083/jcb.200607020.

Derrien, T., R. Johnson, G. Bussotti, A. Tanzer, S. Djebali, H. Tilgner, G. Guernec, D. Martin, A. Merkel, D.G. Knowles, J. Lagarde, L. Veeravalli, X. Ruan, Y. Ruan, T. Lassmann, P. Carninci, J.B. Brown, L. Lipovich, J.M. Gonzalez, M. Thomas, C.A. Davis, R. Shiekhattar, T.R. Gingeras, T.J. Hubbard, C. Notredame, J. Harrow, and R. Guigó. 2012. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 22:1775–1789. doi:10.1101/gr.132159.111.

Deveson, I.W., M.E. Brunck, J. Blackburn, E. Tseng, T. Hon, T.A. Clark, M.B. Clark, J. Crawford, M.E. Dinger, L.K. Nielsen, J.S. Mattick, and T.R. Mercer. 2018. Universal Alternative Splicing of Noncoding Exons. *Cell Syst.* 6:245-255.e5. doi:10.1016/j.cels.2017.12.005.

Dias, A.P., K. Dufu, H. Lei, and R. Reed. 2010. A role for TREX components in the release of spliced mRNA from nuclear speckle domains. *Nat. Commun.* 1:97. doi:10.1038/ncomms1103.

Djebali, S., C.A. Davis, A. Merkel, A. Dobin, T. Lassmann, A. Mortazavi, A. Tanzer, J. Lagarde, W. Lin, F. Schlesinger, C. Xue, G.K. Marinov, J. Khatun, B.A. Williams, C. Zaleski, J. Rozowsky, M. Röder, F. Kokocinski, R.F. Abdelhamid, T. Alioto, I. Antoshechkin, M.T. Baer, N.S. Bar, P. Batut, K. Bell, I. Bell, S. Chakrabortty, X. Chen, J. Chrast, J. Curado, T. Derrien, J. Drenkow, E. Dumais, J. Dumais, R. Duttagupta, E. Falconnet, M. Fastuca, K. Fejes-Toth, P. Ferreira, S. Foissac, M.J. Fullwood, H. Gao, D. Gonzalez, A. Gordon, H. Gunawardena, C. Howald, S. Jha, R. Johnson, P. Kapranov, B. King, C. Kingswood, O.J. Luo, E. Park, K. Persaud, J.B. Preall, P. Ribeca, B. Risk, D. Robyr, M. Sammeth, L. Schaffer, L.-H. See, A. Shahab, J. Skancke, A.M. Suzuki, H. Takahashi, H. Tilgner, D. Trout, N. Walters, H. Wang, J. Wrobel, Y. Yu, X. Ruan, Y. Hayashizaki, J. Harrow, M. Gerstein, T. Hubbard, A. Reymond, S.E. Antonarakis, G. Hannon, M.C. Giddings, Y. Ruan, B. Wold, P. Carninci, R. Guigó, and T.R. Gingeras. 2012. Landscape of transcription in human cells. *Nature*. 489:101–108. doi:10.1038/nature11233.

Dominissini, D., S. Moshitch-Moshkovitz, S. Schwartz, M. Salmon-Divon, L. Ungar, S. Osenberg, K. Cesarkas, J. Jacob-Hirsch, N. Amariglio, M. Kupiec, R. Sorek, and G. Rechavi. 2012. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*. 485:201–206. doi:10.1038/nature11112.

Dominissini, D., S. Nachtergaele, S. Moshitch-Moshkovitz, E. Peer, N. Kol, M.S. Ben-Haim, Q. Dai, A.D. Segni, M. Salmon-Divon, W.C. Clark, G. Zheng, T. Pan, O. Solomon, E. Eyal, V. Hershkovitz, D. Han, L.C. Doré, N. Amariglio, G. Rechavi, and C. He. 2016. The dynamic N^1-

methyladenosine methylome in eukaryotic messenger RNA. *Nature*. 530:441–446. doi:10.1038/nature16998.

Dufu, K., M.J. Livingstone, J. Seebacher, S.P. Gygi, S.A. Wilson, and R. Reed. 2010. ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. *Genes Dev.* 24:2043–2053. doi:10.1101/gad.1898610.

Elbarbary, R.A., W. Li, B. Tian, and L.E. Maquat. 2013. STAU1 binding 3' UTR IRAlus complements nuclear retention to protect cells from PKR-mediated translational shutdown. *Genes Dev.* 27:1495–1510. doi:10.1101/gad.220962.113.

Emerman, M., R. Vazeux, and K. Peden. 1989. The rev gene product of the human immunodeficiency virus affects envelope-specific RNA localization. *Cell*. 57:1155–1165.

Emerson, R.O., and J.H. Thomas. 2009. Adaptive Evolution in Zinc Finger Transcription Factors. *PLOS Genet.* 5:e1000325. doi:10.1371/journal.pgen.1000325.

Erkmann, J.A., R. Sànchez, N. Treichel, W.F. Marzluff, and U. Kutay. 2005. Nuclear export of metazoan replication-dependent histone mRNAs is dependent on RNA length and is mediated by TAP. *RNA N. Y. N.* 11:45–58. doi:10.1261/rna.7189205.

Fan, J., B. Kuai, G. Wu, X. Wu, B. Chi, L. Wang, K. Wang, Z. Shi, H. Zhang, S. Chen, Z. He, S. Wang, Z. Zhou, G. Li, and H. Cheng. 2017. Exosome cofactor hMTR4 competes with export adaptor ALYREF to ensure balanced nuclear RNA pools for degradation and export. *EMBO J.* 36:2870–2886. doi:10.15252/embj.201696139.

Fan, X.C., and J.A. Steitz. 1998. HNS, a nuclear-cytoplasmic shuttling sequence in HuR. *Proc. Natl. Acad. Sci. U. S. A.* 95:15293–15298.

Fleckner, J., M. Zhang, J. Valcárcel, and M.R. Green. 1997. U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. *Genes Dev.* 11:1864–1872.

Fortes, P., Y. Cuevas, F. Guan, P. Liu, S. Pentlicky, S.P. Jung, M.L. Martínez-Chantar, J. Prieto, D. Rowe, and S.I. Gunderson. 2003. Inhibiting expression of specific genes in mammalian cells with 5' end-mutated U1 small nuclear RNAs targeted to terminal exons of pre-mRNA. *Proc. Natl. Acad. Sci. U. S. A.* 100:8264–8269. doi:10.1073/pnas.1332669100.

Fuke, H., and M. Ohno. 2008. Role of poly (A) tail as an identity element for mRNA nuclear export. *Nucleic Acids Res.* 36:1037–1049. doi:10.1093/nar/gkm1120.

Furth, P.A., W.T. Choe, J.H. Rex, J.C. Byrne, and C.C. Baker. 1994. Sequences homologous to 5' splice sites are required for the inhibitory activity of papillomavirus late 3' untranslated regions. *Mol. Cell. Biol.* 14:5278–5289.

Galy, V., O. Gadal, M. Fromont-Racine, A. Romano, A. Jacquier, and U. Nehrbass. 2004. Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. *Cell*. 116:63–73.

Gebhardt, A., M. Habjan, C. Benda, A. Meiler, D.A. Haas, M.Y. Hein, A. Mann, M. Mann, B. Habermann, and A. Pichlmair. 2015. mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3. *Nat. Commun.* 6:8192. doi:10.1038/ncomms9192.

Gong, C., and L.E. Maquat. 2011. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature*. 470:284–288. doi:10.1038/nature09701.

Goraczniak, R., M.A. Behlke, and S.I. Gunderson. 2009. Gene silencing by synthetic U1 adaptors. *Nat. Biotechnol.* 27:257–263. doi:10.1038/nbt.1525.

Gray, M.W., J. Lukeš, J.M. Archibald, P.J. Keeling, and W.F. Doolittle. 2010. Irremediable Complexity? *Science*. 330:920–921. doi:10.1126/science.1198594.

Gregory, T.R. 2005. Synergy between sequence and size in large-scale genomics. *Nat. Rev. Genet.* 6:699–708. doi:10.1038/nrg1674.

Grüter, P., C. Tabernero, C. von Kobbe, C. Schmitt, C. Saavedra, A. Bachi, M. Wilm, B.K. Felber, and E. Izaurralde. 1998. TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Mol. Cell*. 1:649–659. doi:10.1016/S1097-2765(00)80065-9.

Guang, S., and J.E. Mertz. 2005. Pre-mRNA processing enhancer (PPE) elements from intronless genes play additional roles in mRNA biogenesis than do ones from intron-containing genes. *Nucleic Acids Res.* 33:2215–2226. doi:10.1093/nar/gki506.

Gunderson, S.I., M. Polycarpou-Schwarz, and I.W. Mattaj. 1998. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol. Cell*. 1:255–264.

Hacisuleyman, E., L.A. Goff, C. Trapnell, A. Williams, J. Henao-Mejia, L. Sun, P. McClanahan, D.G. Hendrickson, M. Sauvageau, D.R. Kelley, M. Morse, J. Engreitz, E.S. Lander, M. Guttman, H.F. Lodish, R. Flavell, A. Raj, and J.L. Rinn. 2014. Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat. Struct. Mol. Biol.* 21:198–206. doi:10.1038/nsmb.2764.

Hautbergue, G.M., M.-L. Hung, M.J. Walsh, A.P.L. Snijders, C.-T. Chang, R. Jones, C.P. Ponting, M.J. Dickman, and S.A. Wilson. 2009. UIF, a New mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. *Curr. Biol.* 19:1918–1924. doi:10.1016/j.cub.2009.09.041.

Heath, C.G., N. Viphakone, and S.A. Wilson. 2016. The role of TREX in gene expression and disease. *Biochem. J.* 473:2911–2935. doi:10.1042/BCJ20160010.

Hett, A., and S. West. 2014. Inhibition of U4 snRNA in Human Cells Causes the Stable Retention of Polyadenylated Pre-mRNA in the Nucleus. *PLoS ONE*. 9. doi:10.1371/journal.pone.0096174.

Huang, C., D. Liang, D.C. Tatomer, and J.E. Wilusz. 2018. A length-dependent evolutionarily conserved pathway controls nuclear export of circular RNAs. *Genes Dev.* 32:639–644. doi:10.1101/gad.314856.118.

Huang, Y., R. Gattoni, J. Stévenin, and J.A. Steitz. 2003. SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Mol. Cell*. 11:837–843.

Ilan, L., F. Osman, L.S. Namer, E. Eliahu, S. Cohen-Chalamish, Y. Ben-Asouli, Y. Banai, and R. Kaempfer. 2017. PKR activation and eIF2α phosphorylation mediate human globin mRNA splicing at spliceosome assembly. *Cell Res.* 27:688–704. doi:10.1038/cr.2017.39.

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

Jandura, A., and H.M. Krause. 2017. The New RNA World: Growing Evidence for Long Noncoding RNA Functionality. *Trends Genet.* 33:665–676. doi:10.1016/j.tig.2017.08.002.

Johnson, S.A., G. Cubberley, and D.L. Bentley. 2009. Cotranscriptional recruitment of the mRNA export factor Yra1 by direct interaction with the 3' end processing factor Pcf11. *Mol. Cell*. 33:215–226. doi:10.1016/j.molcel.2008.12.007.

Kaessmann, H. 2010. Origins, evolution, and phenotypic impact of new genes. *Genome Res.* 20:1313–1326. doi:10.1101/gr.101386.109.

Kaewsapsak, P., D.M. Shechner, W. Mallard, J.L. Rinn, and A.Y. Ting. 2017. Live-cell mapping of organelle-associated RNAs via proximity biotinylation combined with protein-RNA crosslinking. *eLife*. 6. doi:10.7554/eLife.29224.

Kaida, D., M.G. Berg, I. Younis, M. Kasim, L.N. Singh, L. Wan, and G. Dreyfuss. 2010. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature*. 468:664–668. doi:10.1038/nature09479.

Kaida, D., H. Motoyoshi, E. Tashiro, T. Nojima, M. Hagiwara, K. Ishigami, H. Watanabe, T. Kitahara, T. Yoshida, H. Nakajima, T. Tani, S. Horinouchi, and M. Yoshida. 2007. Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat. Chem. Biol.* 3:576–583. doi:10.1038/nchembio.2007.18.

Kapadia, F., A. Pryor, T.-H. Chang, and L.F. Johnson. 2006. Nuclear localization of poly(A)+ mRNA following siRNA reduction of expression of the mammalian RNA helicases UAP56 and URH49. *Gene*. 384:37–44. doi:10.1016/j.gene.2006.07.010.

Karlin, S., and J. Mrázek. 1997. Compositional differences within and between eukaryotic genomes. *Proc. Natl. Acad. Sci. U. S. A.* 94:10227–10232.

Katahira, J. 2012. mRNA export and the TREX complex. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* 1819:507–513. doi:10.1016/j.bbagrm.2011.12.001.

Katahira, J., D. Okuzaki, H. Inoue, Y. Yoneda, K. Maehara, and Y. Ohkawa. 2013. Human TREX component Thoc5 affects alternative polyadenylation site choice by recruiting mammalian cleavage factor I. *Nucleic Acids Res.* 41:7060–7072. doi:10.1093/nar/gkt414.

Katahira, J., K. Strässer, A. Podtelejnikov, M. Mann, J.U. Jung, and E. Hurt. 1999. The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *EMBO J.* 18:2593–2609. doi:10.1093/emboj/18.9.2593.

Kelley, D., and J. Rinn. 2012. Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biol.* 13:R107. doi:10.1186/gb-2012-13-11-r107.

Kim, D.D.Y., T.T.Y. Kim, T. Walsh, Y. Kobayashi, T.C. Matise, S. Buyske, and A. Gabriel. 2004. Widespread RNA Editing of Embedded Alu Elements in the Human Transcriptome. *Genome Res.* 14:1719–1725. doi:10.1101/gr.2855504.

Kim, J.H., B.B. Lee, Y.M. Oh, C. Zhu, L.M. Steinmetz, Y. Lee, W.K. Kim, S.B. Lee, S. Buratowski, and T. Kim. 2016. Modulation of mRNA and lncRNA expression dynamics by the Set2–Rpd3S pathway. *Nat. Commun.* 7. doi:10.1038/ncomms13534.

de Koning, A.P.J., W. Gu, T.A. Castoe, M.A. Batzer, and D.D. Pollock. 2011. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.* 7:e1002384. doi:10.1371/journal.pgen.1002384.

Kumar, M., and G.G. Carmichael. 1997. Nuclear antisense RNA induces extensive adenosine modifications and nuclear retention of target transcripts. *Proc. Natl. Acad. Sci.* 94:3542–3547. doi:10.1073/pnas.94.8.3542.

Langemeier, J., E.-M. Schrom, A. Rabner, M. Radtke, D. Zychlinski, A. Saborowski, G. Bohn, Y. Mandel-Gutfreund, J. Bodem, C. Klein, and J. Bohne. 2012. A complex immunodeficiency is based on U1 snRNP-mediated poly(A) site suppression. *EMBO J.* 31:4035–4044. doi:10.1038/emboj.2012.252.

Larson, D.R. 2011. What do expression dynamics tell us about the mechanism of transcription? *Curr. Opin. Genet. Dev.* 21:591–599. doi:10.1016/j.gde.2011.07.010.

Le Hir, H., D. Gatfield, E. Izaurralde, and M.J. Moore. 2001. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.* 20:4987–4997. doi:10.1093/emboj/20.17.4987.

Lee, E.S., A. Akef, K. Mahadevan, and A.F. Palazzo. 2015. The consensus 5' splice site motif inhibits mRNA nuclear export. *PLoS One*. 10:e0122743. doi:10.1371/journal.pone.0122743.

Legrain, P., and M. Rosbash. 1989. Some cis- and trans-acting mutants for splicing target pre-mRNA to the cytoplasm. *Cell*. 57:573–583.

Lei, H., A.P. Dias, and R. Reed. 2011. Export and stability of naturally intronless mRNAs require specific coding region sequences and the TREX mRNA export complex. *Proc. Natl. Acad. Sci. U. S. A.* 108:17985–17990. doi:10.1073/pnas.1113076108.

Lei, H., B. Zhai, S. Yin, S. Gygi, and R. Reed. 2013. Evidence that a consensus element found in naturally intronless mRNAs promotes mRNA export. *Nucleic Acids Res.* 41:2517–2525. doi:10.1093/nar/gks1314.

Levanon, E.Y., E. Eisenberg, R. Yelin, S. Nemzer, M. Hallegger, R. Shemesh, Z.Y. Fligelman, A. Shoshan, S.R. Pollock, D. Sztybel, M. Olshansky, G. Rechavi, and M.F. Jantsch. 2004. Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat. Biotechnol.* 22:1001–1005. doi:10.1038/nbt996.

Lévesque, K., M. Halvorsen, L. Abrahamyan, L. Chatel-Chaix, V. Poupon, H. Gordon, L. DesGroseillers, A. Gatignol, and A.J. Mouland. 2006. Trafficking of HIV-1 RNA is mediated by heterogeneous nuclear ribonucleoprotein A2 expression and impacts on viral assembly. *Traffic Cph. Den.* 7:1177–1193.

Li, X., X. Xiong, K. Wang, L. Wang, X. Shu, S. Ma, and C. Yi. 2016a. Transcriptome-wide mapping reveals reversible and dynamic N^1-methyladenosine methylome. *Nat. Chem. Biol.* 12:311–316. doi:10.1038/nchembio.2040.

Li, X., P. Zhu, S. Ma, J. Song, J. Bai, F. Sun, and C. Yi. 2015. Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat. Chem. Biol.* 11:592–597. doi:10.1038/nchembio.1836.

Li, Y., Y. Bor, M.P. Fitzgerald, K.S. Lee, D. Rekosh, and M.-L. Hammarskjold. 2016b. An NXF1 mRNA with a retained intron is expressed in hippocampal and neocortical neurons and is translated into a protein that functions as an Nxf1 cofactor. *Mol. Biol. Cell*. 27:3903–3912. doi:10.1091/mbc.E16-07-0515.

Li, Y., Y.-C. Bor, Y. Misawa, Y. Xue, D. Rekosh, and M.-L. Hammarskjöld. 2006. An intron with a constitutive transport element is retained in a Tap messenger RNA. *Nature*. 443:234–237. doi:10.1038/nature05107.

Lindahl, T. 1993. Instability and decay of the primary structure of DNA. *Nature*. 362:709–715. doi:10.1038/362709a0.

Lu, S., and B.R. Cullen. 2003. Analysis of the stimulatory effect of splicing on mRNA production and utilization in mammalian cells. *RNA*. 9:618–630.

Lu, X.B., J. Heimer, D. Rekosh, and M.L. Hammarskjöld. 1990. U1 small nuclear RNA plays a direct role in the formation of a rev-regulated human immunodeficiency virus env mRNA that remains unspliced. *Proc. Natl. Acad. Sci. U. S. A.* 87:7598–7602.

Lubelsky, Y., and I. Ulitsky. 2018. Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature*. 555:107–111. doi:10.1038/nature25757.

Lucas, B.A., E. Lavi, L. Shiue, H. Cho, S. Katzman, K. Miyoshi, M.C. Siomi, L. Carmel, M. Ares, and L.E. Maquat. 2018. Evidence for convergent evolution of SINE-directed Staufen-mediated mRNA decay. *Proc. Natl. Acad. Sci.* 201715531. doi:10.1073/pnas.1715531115.

Lukeš, J., J.M. Archibald, P.J. Keeling, W.F. Doolittle, and M.W. Gray. 2011. How a neutral evolutionary ratchet can build cellular complexity. *IUBMB Life*. 63:528–537. doi:10.1002/iub.489.

Lund, M.K., and C. Guthrie. 2005. The DEAD-box protein Dbp5p is required to dissociate Mex67p from exported mRNPs at the nuclear rim. *Mol. Cell*. 20:645–651. doi:10.1016/j.molcel.2005.10.005.

Luo, M.J., and R. Reed. 1999. Splicing is required for rapid and efficient mRNA export in metazoans. *Proc. Natl. Acad. Sci. U. S. A.* 96:14937–14942.

Luo, M.L., Z. Zhou, K. Magni, C. Christoforides, J. Rappsilber, M. Mann, and R. Reed. 2001. Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature*. 413:644–647. doi:10.1038/35098106.

Mahboubi, H., E. Seganathy, D. Kong, and U. Stochaj. 2013. Identification of Novel Stress Granule Components That Are Involved in Nuclear Transport. *PLoS ONE*. 8. doi:10.1371/journal.pone.0068356.

Maniatis, T., and R. Reed. 2002. An extensive network of coupling among gene expression machines. *Nature*. 416:499–506. doi:10.1038/416499a.

Martin, W., and E.V. Koonin. 2006. Introns and the origin of nucleus-cytosol compartmentalization. *Nature*. 440:41–45. doi:10.1038/nature04531.

Masuda, S., R. Das, H. Cheng, E. Hurt, N. Dorman, and R. Reed. 2005. Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev.* 19:1512–1517. doi:10.1101/gad.1302205.

Masuyama, K., I. Taniguchi, N. Kataoka, and M. Ohno. 2004. RNA length defines RNA export pathway. *Genes Dev.* 18:2074–2085. doi:10.1101/gad.1216204.

Mauger, O., F. Lemoine, and P. Scheiffele. 2016. Targeted Intron Retention and Excision for Rapid Gene Regulation in Response to Neuronal Activity. *Neuron*. 92:1266–1278. doi:10.1016/j.neuron.2016.11.032.

McCloskey, A., I. Taniguchi, K. Shinmyozu, and M. Ohno. 2012. hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. *Science*. 335:1643–1646. doi:10.1126/science.1218469.

Melé, M., K. Mattioli, W. Mallard, D.M. Shechner, C. Gerhardinger, and J.L. Rinn. 2017. Chromatin environment, transcriptional regulation, and splicing distinguish lincRNAs and mRNAs. *Genome Res.* 27:27–37. doi:10.1101/gr.214205.116.

Meola, N., M. Domanski, E. Karadoulama, Y. Chen, C. Gentil, D. Pultz, K. Vitting-Seerup, S. Lykke-Andersen, J.S. Andersen, A. Sandelin, and T.H. Jensen. 2016. Identification of a Nuclear Exosome Decay Pathway for Processed Transcripts. *Mol. Cell*. 64:520–533. doi:10.1016/j.molcel.2016.09.025.

Meyer, K.D., Y. Saletore, P. Zumbo, O. Elemento, C.E. Mason, and S.R. Jaffrey. 2012. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell*. 149:1635–1646. doi:10.1016/j.cell.2012.05.003.

Miyagawa, R., K. Tano, R. Mizuno, Y. Nakamura, K. Ijiri, R. Rakwal, J. Shibato, Y. Masuo, A. Mayeda, T. Hirose, and N. Akimitsu. 2012. Identification of cis- and trans-acting factors involved in the localization of MALAT-1 noncoding RNA to nuclear speckles. *RNA*. 18:738–751. doi:10.1261/rna.028639.111.

Moore, M.J., and N.J. Proudfoot. 2009. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell*. 136:688–700. doi:10.1016/j.cell.2009.02.001.

Mukherjee, N., L. Calviello, A. Hirsekorn, S. de Pretis, M. Pelizzola, and U. Ohler. 2017. Integrative classification of human coding and noncoding genes through RNA metabolism profiles. *Nat. Struct. Mol. Biol.* 24:86–96. doi:10.1038/nsmb.3325.

Müller-McNicoll, M., V. Botti, A.M. de Jesus Domingues, H. Brandl, O.D. Schwich, M.C. Steiner, T. Curk, I. Poser, K. Zarnack, and K.M. Neugebauer. 2016. SR proteins are NXF1 adaptors that link alternative RNA processing to mRNA export. *Genes Dev.* 30:553–566. doi:10.1101/gad.276477.115.

Naro, C., A. Jolly, S. Di Persio, P. Bielli, N. Setterblad, A.J. Alberdi, E. Vicini, R. Geremia, P. De la Grange, and C. Sette. 2017. An Orchestrated Intron Retention Program in Meiosis Controls Timely Usage of Transcripts during Germ Cell Differentiation. *Dev. Cell*. 41:82-93.e4. doi:10.1016/j.devcel.2017.03.003.

Nasioulas, G., A.S. Zolotukhin, C. Tabernero, L. Solomin, C.P. Cunningham, G.N. Pavlakis, and B.K. Felber. 1994. Elements distinct from human immunodeficiency virus type 1 splice sites are responsible for the Rev dependence of env mRNA. *J. Virol.* 68:2986–2993.

Noh, J.H., K.M. Kim, K. Abdelmohsen, J.-H. Yoon, A.C. Panda, R. Munk, J. Kim, J. Curtis, C.A. Moad, C.M. Wohler, F.E. Indig, W. de Paula, D.B. Dudekula, S. De, Y. Piao, X. Yang, J.L. Martindale, R. de Cabo, and M. Gorospe. 2016. HuR and GRSF1 modulate the nuclear export and mitochondrial localization of the lncRNA RMRP. *Genes Dev.* 30:1224–1239. doi:10.1101/gad.276022.115.

Nojima, T., T. Hirose, H. Kimura, and M. Hagiwara. 2007. The Interaction between Cap-binding Complex and RNA Export Factor Is Required for Intronless mRNA Export. *J. Biol. Chem.* 282:15645–15651. doi:10.1074/jbc.M700629200.

Nott, A., H. Le Hir, and M.J. Moore. 2004. Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. *Genes Dev.* 18:210–222. doi:10.1101/gad.1163204.

Ogami, K., P. Richard, Y. Chen, M. Hoque, W. Li, J.J. Moresco, J.R. Yates, B. Tian, and J.L. Manley. 2017. An Mtr4/ZFC3H1 complex facilitates turnover of unstable nuclear RNAs to prevent their cytoplasmic transport and global translational repression. *Genes Dev.* doi:10.1101/gad.302604.117.

Palazzo, A.F., and A. Akef. 2012. Nuclear export as a key arbiter of "mRNA identity" in eukaryotes. *Biochim. Biophys. Acta*. 1819:566–577. doi:10.1016/j.bbagrm.2011.12.012.

Palazzo, A.F., and T.R. Gregory. 2014. The case for junk DNA. *PLoS Genet.* 10:e1004351. doi:10.1371/journal.pgen.1004351.

Palazzo, A.F., and E.S. Lee. 2015. Non-coding RNA: what is functional and what is junk? *Front. Genet.* 6:2. doi:10.3389/fgene.2015.00002.

Palazzo, A.F., M. Springer, Y. Shibata, C.-S. Lee, A.P. Dias, and T.A. Rapoport. 2007. The signal sequence coding region promotes nuclear export of mRNA. *PLoS Biol.* 5:e322. doi:10.1371/journal.pbio.0050322.

Palazzo, A.F., and M. Truong. 2016. Single particle imaging of mRNAs crossing the nuclear pore: Surfing on the edge. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 38:744–750. doi:10.1002/bies.201600038.

Park, E., and L.E. Maquat. 2013. Staufen-mediated mRNA decay. *Wiley Interdiscip. Rev. RNA*. 4:423–435. doi:10.1002/wrna.1168.

Polson, A.G., P.F. Crain, S.C. Pomerantz, J.A. McCloskey, and B.L. Bass. 1991. The mechanism of adenosine to inosine conversion by the double-stranded RNA unwinding/modifying activity: a high-performance liquid chromatography-mass spectrometry analysis. *Biochemistry*. 30:11507–11514.

Prasanth, K.V., S.G. Prasanth, Z. Xuan, S. Hearn, S.M. Freier, C.F. Bennett, M.Q. Zhang, and D.L. Spector. 2005. Regulating Gene Expression through RNA Nuclear Retention. *Cell*. 123:249–263. doi:10.1016/j.cell.2005.08.033.

Proudfoot, N.J. 2016. Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science*. 352:aad9926. doi:10.1126/science.aad9926.

Rain, J.-C., and P. Legrain. 1997. In vivo commitment to splicing in yeast involves the nucleotide upstream from the branch site conserved sequence and the Mud2 protein. *EMBO J.* 16:1759–1771. doi:10.1093/emboj/16.7.1759.

Rajanala, K., and V.K. Nandicoori. 2012. Localization of nucleoporin Tpr to the nuclear pore complex is essential for Tpr mediated regulation of the export of unspliced RNA. *PLoS One*. 7:e29921. doi:10.1371/journal.pone.0029921.

Reese, M.G., F.H. Eeckman, D. Kulp, and D. Haussler. 1997. Improved splice site detection in Genie. *J. Comput. Biol. J. Comput. Mol. Cell Biol.* 4:311–323. doi:10.1089/cmb.1997.4.311.

Roundtree, I.A., G.-Z. Luo, Z. Zhang, X. Wang, T. Zhou, Y. Cui, J. Sha, X. Huang, L. Guerrero, P. Xie, E. He, B. Shen, and C. He. 2017. YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *eLife*. 6. doi:10.7554/eLife.31311.

Rowe, H.M., and D. Trono. 2011. Dynamic control of endogenous retroviruses during development. *Virology*. 411:273–287. doi:10.1016/j.virol.2010.12.007.

Roy, D., J. Bhanja Chowdhury, and S. Ghosh. 2013. Polypyrimidine tract binding protein (PTB) associates with intronic and exonic domains to squelch nuclear export of unspliced RNA. *FEBS Lett.* 587:3802–3807. doi:10.1016/j.febslet.2013.10.005.

Ruepp, M.-D., C. Aringhieri, S. Vivarelli, S. Cardinale, S. Paro, D. Schümperli, and S.M.L. Barabino. 2009. Mammalian pre-mRNA 3′ End Processing Factor CF Im68 Functions in mRNA Export. *Mol. Biol. Cell*. 20:5211–5223. doi:10.1091/mbc.E09-05-0389.

Schilders, G., E. van Dijk, and G.J.M. Pruijn. 2007. C1D and hMtr4p associate with the human exosome subunit PM/Scl-100 and are involved in pre-rRNA processing. *Nucleic Acids Res.* 35:2564–2572. doi:10.1093/nar/gkm082.

Schmitges, F.W., E. Radovani, H.S. Najafabadi, M. Barazandeh, L.F. Campitelli, Y. Yin, A. Jolma, G. Zhong, H. Guo, T. Kanagalingam, W.F. Dai, J. Taipale, A. Emili, J.F. Greenblatt, and T.R. Hughes. 2016. Multiparameter functional diversity of human C2H2 zinc finger proteins. *Genome Res.* 26:1742–1752. doi:10.1101/gr.209643.116.

Schwartz, S., D.A. Bernstein, M.R. Mumbach, M. Jovanovic, R.H. Herbst, B.X. León-Ricardo, J.M. Engreitz, M. Guttman, R. Satija, E.S. Lander, G. Fink, and A. Regev. 2014. Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell*. 159:148–162. doi:10.1016/j.cell.2014.08.028.

Séguin, B., A. Staffa, and A. Cochrane. 1998. Control of Human Immunodeficiency Virus Type 1 RNA Metabolism: Role of Splice Sites and Intron Sequences in Unspliced Viral RNA Subcellular Distribution. *J. Virol.* 72:9503–9513.

Shi, M., H. Zhang, X. Wu, Z. He, L. Wang, S. Yin, B. Tian, G. Li, and H. Cheng. 2017. ALYREF mainly binds to the 5' and the 3' regions of the mRNA in vivo. *Nucleic Acids Res.* 45:9640–9653. doi:10.1093/nar/gkx597.

Shibata, S., Y. Matsuoka, and Y. Yoneda. 2002. Nucleocytoplasmic transport of proteins and poly(A)+ RNA in reconstituted Tpr-less nuclei in living mammalian cells. *Genes Cells*. 7:421–434.

Shukla, C.J., A.L. McCorkindale, C. Gerhardinger, K.D. Korthauer, M.N. Cabili, D.M. Shechner, R.A. Irizarry, P.G. Maass, and J.L. Rinn. 2018. High-throughput identification of RNA nuclear enrichment sequences. *EMBO J.* 37. doi:10.15252/embj.201798452.

Singh, G., A. Kucukural, C. Cenik, J.D. Leszyk, S.A. Shaffer, Z. Weng, and M.J. Moore. 2012. The Cellular EJC Interactome Reveals Higher-Order mRNP Structure and an EJC-SR Protein Nexus. *Cell*. 151:915–916. doi:10.1016/j.cell.2012.10.007.

Squires, J.E., H.R. Patel, M. Nousch, T. Sibbritt, D.T. Humphreys, B.J. Parker, C.M. Suter, and T. Preiss. 2012. Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res.* 40:5023–5033. doi:10.1093/nar/gks144.

Stoltzfus, A. 1999. On the possibility of constructive neutral evolution. *J. Mol. Evol.* 49:169–181.

Stoltzfus, A. 2012. Constructive neutral evolution: exploring evolutionary theory's curious disconnect. *Biol. Direct*. 7:35. doi:10.1186/1745-6150-7-35.

Strässer, K., and E. Hurt. 2001. Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. *Nature*. 413:648–652. doi:10.1038/35098113.

Strässer, K., S. Masuda, P. Mason, J. Pfannstiel, M. Oppizzi, S. Rodriguez-Navarro, A.G. Rondón, A. Aguilera, K. Struhl, R. Reed, and E. Hurt. 2002. TREX is a conserved complex coupling transcription with messenger RNA export. *Nature*. 417:304–308. doi:10.1038/nature746.

Takata, M.A., D. Gonçalves-Carneiro, T.M. Zang, S.J. Soll, A. York, D. Blanco-Melo, and P.D. Bieniasz. 2017. CG dinucleotide suppression enables antiviral defence targeting non-self RNA. *Nature*. 550:124–127. doi:10.1038/nature24039.

Takemura, R., T. Takeiwa, I. Taniguchi, A. McCloskey, and M. Ohno. 2011. Multiple factors in the early splicing complex are involved in the nuclear retention of pre-mRNAs in mammalian cells. *Genes Cells*. 16:1035–1049. doi:10.1111/j.1365-2443.2011.01548.x.

Tan, W., M. Schalling, C. Zhao, M. Luukkonen, M. Nilsson, E.M. Fenyö, G.N. Pavlakis, and S. Schwartz. 1996. Inhibitory activity of the equine infectious anemia virus major 5' splice site in the absence of Rev. *J. Virol.* 70:3645–3658.

Taniguchi, I., and M. Ohno. 2008. ATP-dependent recruitment of export factor Aly/REF onto intronless mRNAs by RNA helicase UAP56. *Mol. Cell. Biol.* 28:601–608. doi:10.1128/MCB.01341-07.

Tilgner, H., D.G. Knowles, R. Johnson, C.A. Davis, S. Chakrabortty, S. Djebali, J. Curado, M. Snyder, T.R. Gingeras, and R. Guigó. 2012. Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res.* 22:1616–1625. doi:10.1101/gr.134445.111.

Topisirovic, I., N. Siddiqui, V.L. Lapointe, M. Trost, P. Thibault, C. Bangeranye, S. Piñol-Roma, and K.L.B. Borden. 2009. Molecular dissection of the eukaryotic initiation factor 4E (eIF4E) export-competent RNP. *EMBO J.* 28:1087–1098. doi:10.1038/emboj.2009.53.

Tran, D.D.H., S. Saran, A.J.K. Williamson, A. Pierce, O. Dittrich-Breiholz, L. Wiehlmann, A. Koch, A.D. Whetton, and T. Tamura. 2014. THOC5 controls 3′end-processing of immediate early genes via interaction with polyadenylation specific factor 100 (CPSF100). *Nucleic Acids Res.* 42:12249–12260. doi:10.1093/nar/gku911.

Tress, M.L., F. Abascal, and A. Valencia. 2017. Alternative Splicing May Not Be the Key to Proteome Complexity. *Trends Biochem. Sci.* 42:98–110. doi:10.1016/j.tibs.2016.08.008.

Umlauf, D., J. Bonnet, F. Waharte, M. Fournier, M. Stierle, B. Fischer, L. Brino, D. Devys, and L. Tora. 2013. The human TREX-2 complex is stably associated with the nuclear pore basket. *J. Cell Sci.* 126:2656–2667. doi:10.1242/jcs.118000.

Vagner, S., U. Rüegsegger, S.I. Gunderson, W. Keller, and I.W. Mattaj. 2000. Position-dependent inhibition of the cleavage step of pre-mRNA 3'-end processing by U1 snRNP. *RNA N. Y. N.* 6:178–188.

Valencia, P., A.P. Dias, and R. Reed. 2008. Splicing promotes rapid and efficient mRNA export in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 105:3386–3391. doi:10.1073/pnas.0800250105.

Vargas, D.Y., K. Shah, M. Batish, M. Levandoski, S. Sinha, S.A.E. Marras, P. Schedl, and S. Tyagi. 2011. Single-molecule imaging of transcriptionally coupled and uncoupled splicing. *Cell*. 147:1054–1065. doi:10.1016/j.cell.2011.10.024.

Vinciguerra, P., N. Iglesias, J. Camblong, D. Zenklusen, and F. Stutz. 2005. Perinuclear Mlp proteins downregulate gene expression in response to a defect in mRNA export. *EMBO J.* 24:813–823. doi:10.1038/sj.emboj.7600527.

Wang, B., D. Rekosh, and M.-L. Hammarskjold. 2015. Evolutionary conservation of a molecular machinery for export and expression of mRNAs with retained introns. *RNA*. 21:426–437. doi:10.1261/rna.048520.114.

Weirich, C.S., J.P. Erzberger, J.S. Flick, J.M. Berger, J. Thorner, and K. Weis. 2006. Activation of the DExD/H-box protein Dbp5 by the nuclear-pore protein Gle1 and its coactivator InsP6 is required for mRNA export. *Nat. Cell Biol.* 8:668–676. doi:10.1038/ncb1424.

van Werven, F.J., G. Neuert, N. Hendrick, A. Lardenois, S. Buratowski, A. van Oudenaarden, M. Primig, and A. Amon. 2012. Transcription of two long non-coding RNAs mediates mating type control of gametogenesis in budding yeast. *Cell*. 150:1170–1181. doi:10.1016/j.cell.2012.06.049.

Wickramasinghe, V.O., R. Andrews, P. Ellis, C. Langford, J.B. Gurdon, M. Stewart, A.R. Venkitaraman, and R.A. Laskey. 2014. Selective nuclear export of specific classes of mRNA from mammalian nuclei is promoted by GANP. *Nucleic Acids Res.* 42:5059–71. doi:10.1093/nar/gku095.

Wickramasinghe, V.O., P.I.A. McMurtrie, A.D. Mills, Y. Takei, S. Penrhyn-Lowe, Y. Amagase, S. Main, J. Marr, M. Stewart, and R.A. Laskey. 2010. mRNA export from mammalian cell nuclei is dependent on GANP. *Curr. Biol.* 20:25–31. doi:10.1016/j.cub.2009.10.078.

Wilk, R., J. Hu, D. Blotsky, and H.M. Krause. 2016. Diverse and pervasive subcellular distributions for both coding and long noncoding RNAs. *Genes Dev.* 30:594–609. doi:10.1101/gad.276931.115.

Wilusz, J.E., C.K. JnBaptiste, L.Y. Lu, C.-D. Kuhn, L. Joshua-Tor, and P.A. Sharp. 2012. A triple helix stabilizes the 3′ ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev.* 26:2392–2407. doi:10.1101/gad.204438.112.

Woo, H., S. Dam Ha, S.B. Lee, S. Buratowski, and T. Kim. 2017. Modulation of gene expression dynamics by co-transcriptional histone methylations. *Exp. Mol. Med.* 49:e326. doi:10.1038/emm.2017.19.

Yang, X., Y. Yang, B.-F. Sun, Y.-S. Chen, J.-W. Xu, W.-Y. Lai, A. Li, X. Wang, D.P. Bhattarai, W. Xiao, H.-Y. Sun, Q. Zhu, H.-L. Ma, S. Adhikari, M. Sun, Y.-J. Hao, B. Zhang, C.-M. Huang, N. Huang, G.-B. Jiang, Y.-L. Zhao, H.-L. Wang, Y.-P. Sun, and Y.-G. Yang. 2017. 5-methylcytosine promotes mRNA export — NSUN2 as the methyltransferase and ALYREF as an m5C reader. *Cell Res.* 27:606–625. doi:10.1038/cr.2017.55.

Yap, K., Z.Q. Lim, P. Khandelia, B. Friedman, and E.V. Makeyev. 2012. Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. *Genes Dev.* 26:1209–1223. doi:10.1101/gad.188037.112.

Zhang, B., L. Gunawardane, F. Niazi, F. Jahanbani, X. Chen, and S. Valadkhan. 2014a. A novel RNA motif mediates the strict nuclear localization of a long noncoding RNA. *Mol. Cell. Biol.* 34:2318–2329. doi:10.1128/MCB.01673-13.

Zhang, H., K. Mahadevan, and A.F. Palazzo. 2014b. Sumoylation is Required for the Cytoplasmic Accumulation of a Subset of mRNAs. *Genes*. 5:982–1000. doi:10.3390/genes5040982.

Zhang, K., J.G. Daigle, K.M. Cunningham, A.N. Coyne, K. Ruan, J.C. Grima, K.E. Bowen, H. Wadhwa, P. Yang, F. Rigo, J.P. Taylor, A.D. Gitler, J.D. Rothstein, and T.E. Lloyd. 2018. Stress Granule Assembly Disrupts Nucleocytoplasmic Transport. *Cell*. 173:958-971.e17. doi:10.1016/j.cell.2018.03.025.

Zhang, Z., and G.G. Carmichael. 2001. The Fate of dsRNA in the Nucleus. *Cell*. 106:465–476. doi:10.1016/S0092-8674(01)00466-4.

Zheng, G., J.A. Dahl, Y. Niu, P. Fedorcsak, C.-M. Huang, C.J. Li, C.B. Vågbø, Y. Shi, W.-L. Wang, S.-H. Song, Z. Lu, R.P.G. Bosmans, Q. Dai, Y.-J. Hao, X. Yang, W.-M. Zhao, W.-M. Tong, X.-J. Wang, F. Bogdan, K. Furu, Y. Fu, G. Jia, X. Zhao, J. Liu, H.E. Krokan, A. Klungland, Y.-G. Yang, and C. He. 2013. ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. *Mol. Cell*. 49:18–29. doi:10.1016/j.molcel.2012.10.015.

Figure 1. The mRNA nuclear export pathway

The TREX complex, which is composed of the Tho complex, the RNA helicase UAP56 (or its paralog URH49) and Aly are loaded onto the mRNA co-transcriptionally or by processing events. At some point, the UAP56 hydrolyses ATP and then is replaced by the nuclear export receptor composed of Nxf1 and Nxt1 (also known as TAP and p15) to form an export competent mRNP. The Nxf1-Nxt1 heterodimer physically interacts with the FG repeats of Nups to ferry its cargo across the nuclear pore. GANP, which forms part o the TREX2 complex is also required for export, although its exact role is not understood. After passing

through the nuclear pore complex, the mRNP is furthered remodelled by cytoplasmic pore-associated proteins such as Gle1, Dbp5 and Rae1/Gle2. It is though that these remodelling events remove certain nuclear associated exported factors, which are then recycled back into the nuclear pore. In some cases these mRNP remodeling events render the mRNP more 'translationally' competent (Palazzo and Truong, 2016). Factors that are essential for mRNA export are depicted in red, other export factors are depicted in yellow.

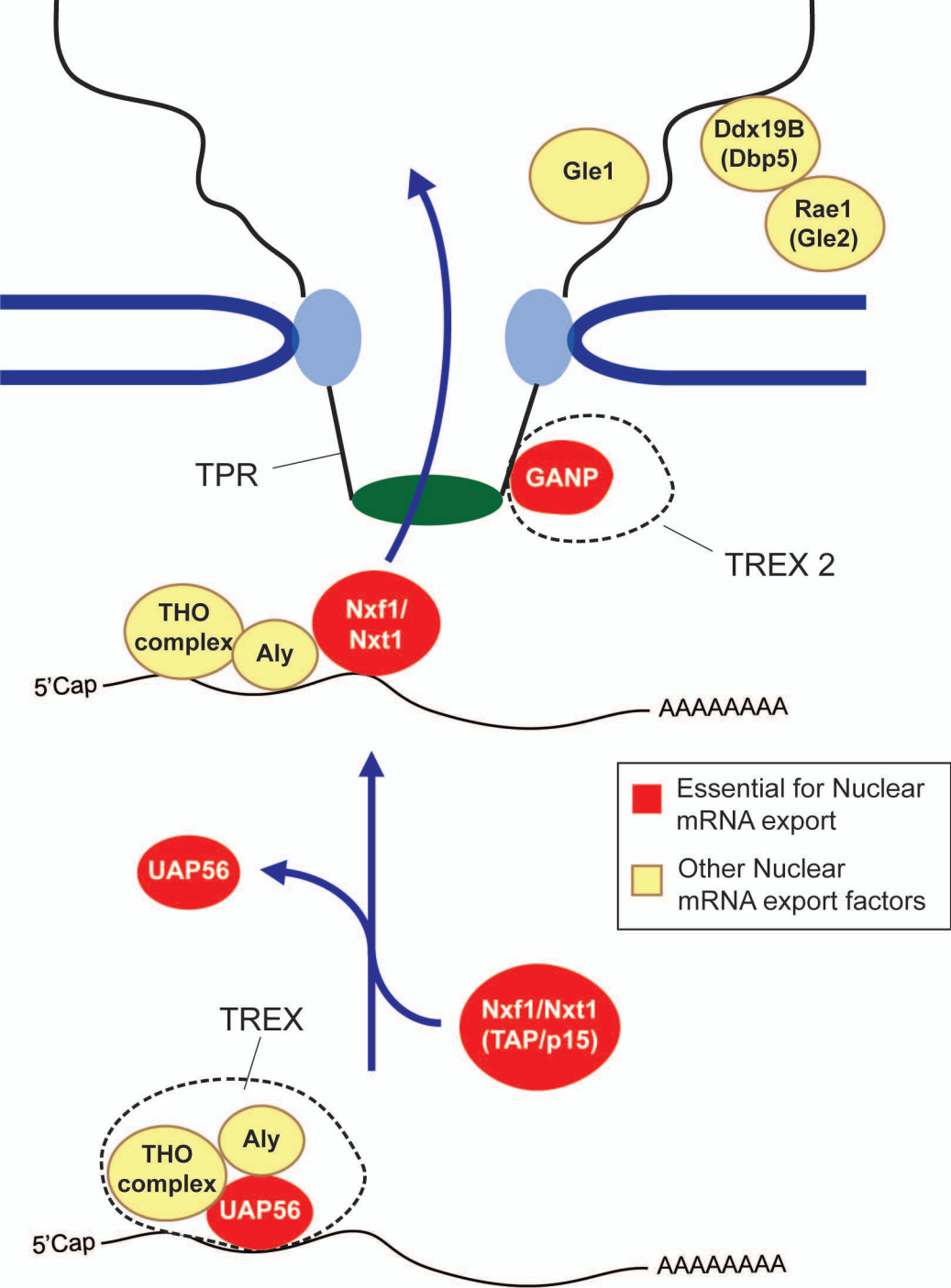
Figure 2. The 5' splice site motif promotes nuclear retention of RNAs

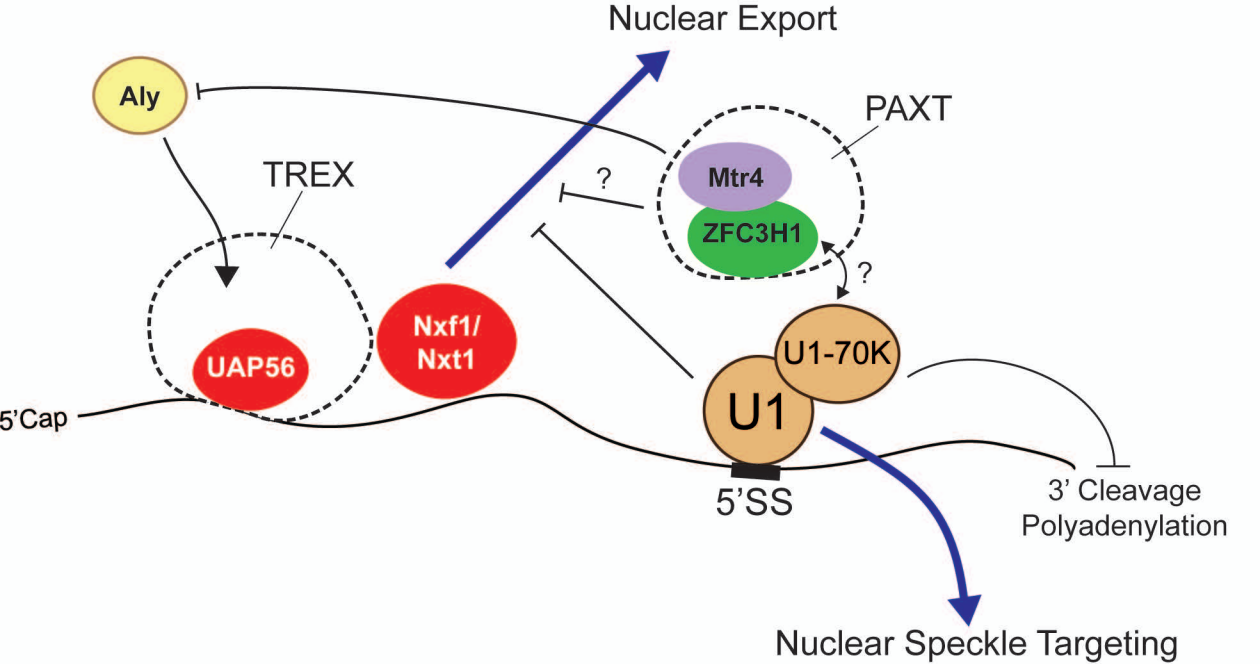
The presence of the 5' splice site (5'SS) motif in the 3' untranslated region (3'UTR) promotes nuclear retention. U1 snRNP recognizes the motif and may recruit nuclear surveillance machinery (e.g. PAXT complex) through U1-70K. Interestingly, although these RNAs are able to recruit UAP56 and Nxf1/Nxt1, they are not exported. Mtr4, a member of the PAXT complex and a co-activator/co-adaptor of the nuclear exosome, competes with Aly for its association with the RNA and various 5' cap binding proteins. Other work suggests that the 5' splice site motif suppresses premature cleavage and polyadenylation through the action of U1-70K.

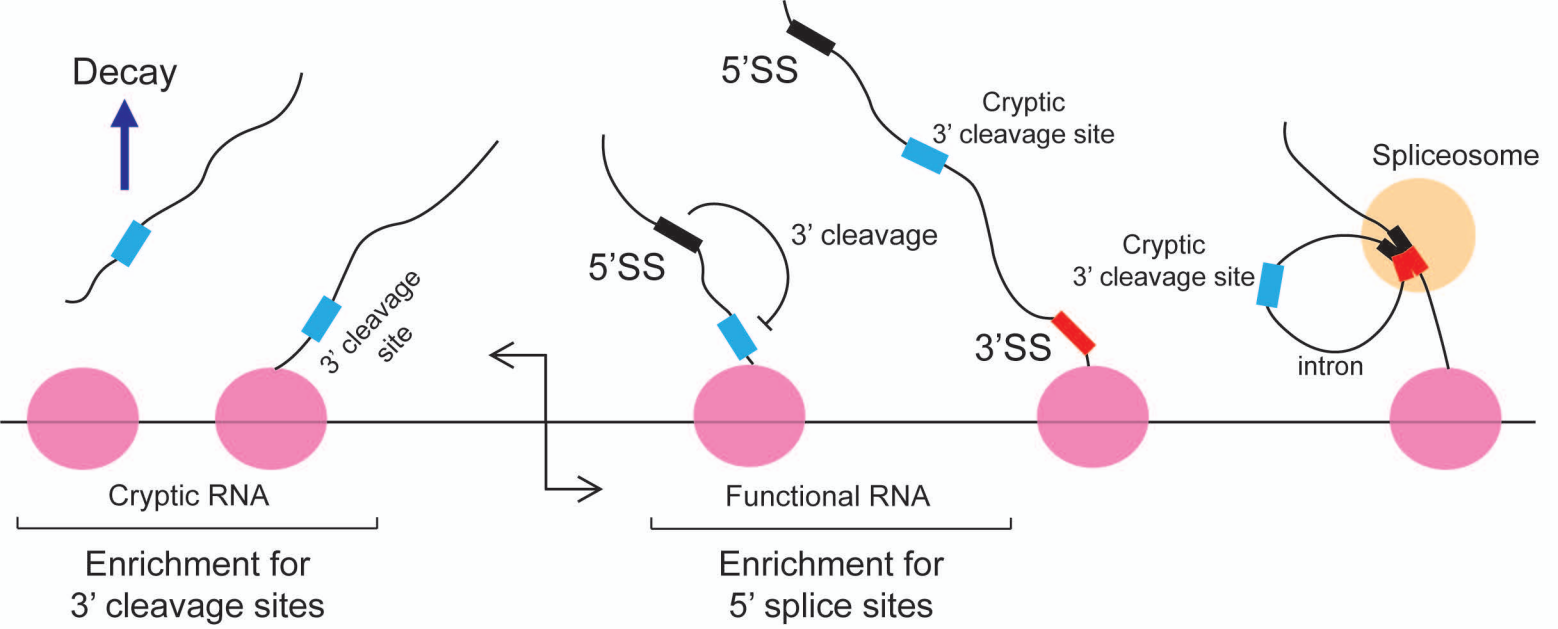
Figure 3. The 5' splice site motif suppresses premature cleavage and polyadenylation

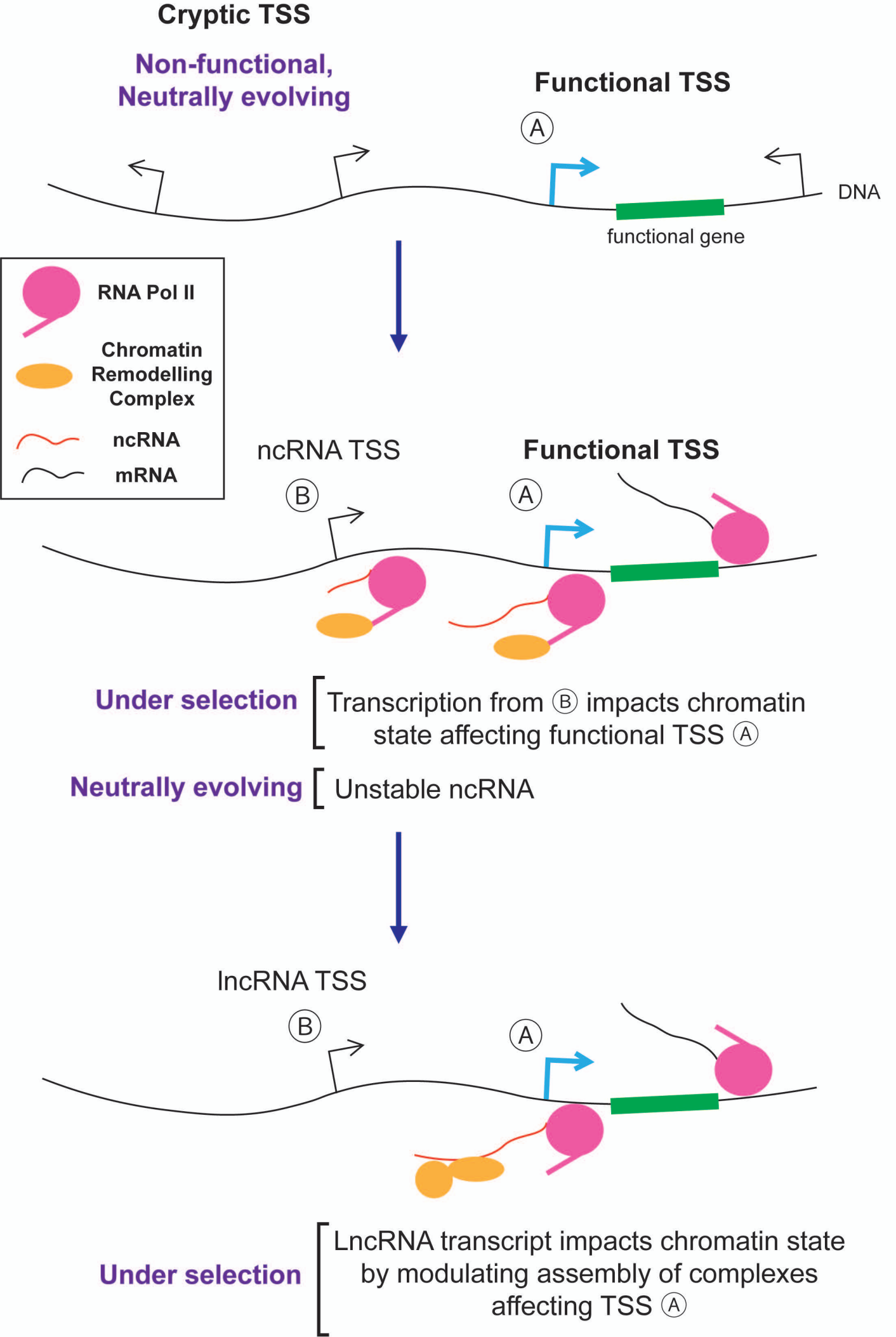
Almada et al., found that in bidirectional promoters that produce one stable transcript, 5'SS motif are enriched in the sense direction (stable RNA) while 3' cleavage sites are enriched in the anti-sense direction (unstable RNA). Under normal circumstances these cryptic unstable RNAs are cleaved and degraded. In the sense direction, the 5' splice site motif suppresses the use of downstream cryptic 3' cleavage sites, allowing RNA PolII to synthesize the RNA transcript without the recruitment of the 3' end processing machinery. These cryptic 3' cleavage sites are typically present in introns and are removed during splicing.

Figure 4. The evolution of lncRNAs from junk RNAs


Gle1
Ddx19B
(Dbp5)
Rae1
(Gle2)
TPR
GANP
TREX 2
THO
complex
Aly
Nxf1/
Nxt1
5'Cap
AAAAAAAA
Essential for Nuclear
mRNA export
Other Nuclear
mRNA export factors
UAP56
Nxf1/Nxt1
(TAP/p15)
TREX
Aly
THO
complex
UAP56
5'Cap
AAAAAAAA


Nuclear Export
Aly
PAXT
TREX
?
Mtr4
ZFC3H1
?
Nxf1/
Nxt1
UAP56
U1-70K
U1
5'Cap
5'SS
3' Cleavage
Polyadenylation
Nuclear Speckle Targeting


Decay
3' cleavage site
Cryptic RNA
Enrichment for
3' cleavage sites
5'SS
3' cleavage
5'SS
Cryptic
3' cleavage site
3'SS
Functional RNA
Enrichment for
5' splice sites
Spliceosome
Cryptic
3' cleavage site
intron


Cryptic TSS
Non-functional,
Neutrally evolving
Functional TSS
A
DNA
functional gene
RNA Pol II
Chromatin
Remodelling
Complex
ncRNA
mRNA
ncRNA TSS
B
Functional TSS
A
Under selection
Transcription from B impacts chromatin
state affecting functional TSS A
Neutrally evolving
Unstable ncRNA
lncRNA TSS
B
A
Under selection
LncRNA transcript impacts chromatin state
by modulating assembly of complexes
affecting TSS A