

REVIEW



mRNA nuclear export: how mRNA identity features distinguish functional RNAs from junk transcripts

Alexander F. Palazzo , Yi Qiu, and Yoon Mo Kang

Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada

ABSTRACT

The division of the cellular space into nucleoplasm and cytoplasm promotes quality control mechanisms that prevent misprocessed mRNAs and junk RNAs from gaining access to the translational machinery. Here, we explore how properly processed mRNAs are distinguished from both misprocessed mRNAs and junk RNAs by the presence or absence of various ‘identity features’.

ARTICLE HISTORY

Revised 27 November 2023
Accepted 5 December 2023

KEYWORDS

mRNA nuclear export;
nuclear retention; splicing;
GC-content; mRNA quality
control

How the nucleocytoplasmic division acts as a quality control system for mRNA metabolism

The eukaryotic cellular space is divided into two regions, the nucleus where precursor messenger RNA (sometimes referred to as premature mRNA and is abbreviated as pre-mRNA) is synthesized and processed to form mature mRNA, and the cytoplasm where the mature mRNA is translated into protein [1,2].

This division has several benefits. First, it allows RNA processing machinery to operate on pre-mRNAs without interference from the translation machinery and vice versa. Second, this division prevents the translation of inappropriately processed pre-mRNAs, which can be deleterious to the cell. Processed transcripts are evaluated in the nucleoplasm by quality control processes and those that are misprocessed are eliminated before they reach the translational machinery in the cytoplasm. Misprocessed transcripts include those that are spliced using suboptimal exon-intron boundaries, and those cleaved pre-maturely by the polyadenylation machinery to generate intronic polyadenylation transcripts. Third, since eukaryotic genomes are largely composed of non-functional DNA that are nevertheless transcribed, the nucleocytoplasmic division allows for quality control machinery to eliminate non-functional RNAs in the nucleus before they have the chance of encountering ribosomes [3–5]. Note that there is still some debate as to whether non-functional, or ‘junk’, RNA exists. Although some long non-coding RNAs are no-doubt functional, even then most optimistic estimates suggest that they are transcribed from no more than 2% of the human genome [6] (including introns this rises to about 10%). In addition, about 35% of the genome is transcribed into pre-mRNA, of which 2% is exonic (i.e. present in the final processed product) [7]. In contrast, greater than 80% of the genome is transcribed at some level in some cell type [8],

and the majority of these loci are non-functional based on both conservation estimates and biochemical data [4,6,9–12]. Thus, despite what a few critics of junk DNA/RNA claim [13,14], most of the data in the literature support the idea that eukaryotes produce a sizable amount of junk RNA. Furthermore, as described in the next section, this view is consistent with the co-evolution of the nucleus, splicing and junk DNA.

The evolution of the nucleocytoplasmic divide and mRNA metabolism

Over the past two decades, it has become clear that the origins of the nucleocytoplasmic divide and the expansion of mRNA metabolism likely co-evolved [1,2]. It is believed that during eukaryogenesis, at least two organisms entered an endosymbiotic relationship, the first was an alpha-proteobacteria, which in time evolved into present-day mitochondria, while the second was an archaeon that itself had acquired a substantial number of eubacterial genes and may have been the product of a prior endosymbiotic event [15]. Over time, genes from the alpha-proteobacteria were absorbed into the archaea genome to form the nuclear genome. This included not only genes that code for mitochondrial-targeted proteins but also Group II introns, which eventually evolved into our spliceosome [16–18], and likely the original introns that were present in the last eukaryote common ancestor (LECA), which appears to have been intron-rich [18,19].

What remains unclear is the exact timing of when the nucleus appeared. The nucleus is a subdomain of the eukaryotic endomembrane system (i.e. the endoplasmic reticulum), and recent analyses suggest that these membranes existed in the archeal branch of our ancestry, prior to either the appearance of a nuclear pore or the acquisition of mitochondria [20].

Despite this, it remains unclear whether the nuclear envelope and nuclear pore complex evolved prior to splicing. Regardless of the exact timeline, once the nucleocytoplasmic divide was created, it likely promoted the proliferation of introns by reducing the deleteriousness of the byproducts of splicing as described above. It also likely promoted the proliferation of intergenic DNA by reducing the deleteriousness of spurious transcription [6,9].

mRNA nuclear export: its primary role in quality control and its co-option to regulate gene expression

It is thus clear that the original, and still primary, role of mRNA export is to sort properly processed mRNAs from both misprocessed mRNAs and junk transcripts [3,5]. The mRNA export machinery accomplishes this by recognizing mRNA identity features, the most important ones being splicing [21,22] and GC-content [3,23–27] (Figure 1A). As a consequence, most mRNAs utilize one of the two main export pathways, the splicing-dependent pathway, and the GC-dependent pathway, which is sometimes referred to as the alternative mRNA export pathway, or ALREX. These same mRNA export pathways are also used by functional long non-coding RNAs that have roles in the cytoplasm. At the same time, eukaryotic cells contain a number of nuclear retention and decay pathways that recognize other features

[5,28]. In some cases, mRNA identity features actively repress these nuclear retention pathways [29], while in other cases nuclear retention pathways actively suppress nuclear export pathways [30,31]. It is also likely that in many cases mRNAs that are targeted for nuclear export, simply evade nuclear RNA decay pathways [28]. Thus together, mRNA nuclear export pathways, mRNA nuclear retention pathways and RNA decay pathways coordinately act to promote the expression of functional RNAs (i.e. mRNAs and functional cytoplasmic lncRNAs) while suppressing the expression of junk transcripts [6].

Although the nucleocytoplasmic divide functions primarily to prevent unspliced pre-mRNAs, misprocessed mRNAs, and junk RNAs from entering into contact with ribosomes, the nuclear export machinery has been co-opted to regulate gene expression. Some of these include the specialized regulation of mRNA export for genes involved in cell cycle progression [32–34], innate immune activity [35], heat shock response [36] and metabolic homeostasis [37]. In some cases, specialized export pathways recognize unique elements only under certain conditions, while in other cases particular ‘detained’ introns remain unspliced and their removal activates mRNA export [38]. The kinetics of mRNA nuclear export also helps to dampen fluctuations in protein levels that would otherwise occur due to bursts in transcription [39].

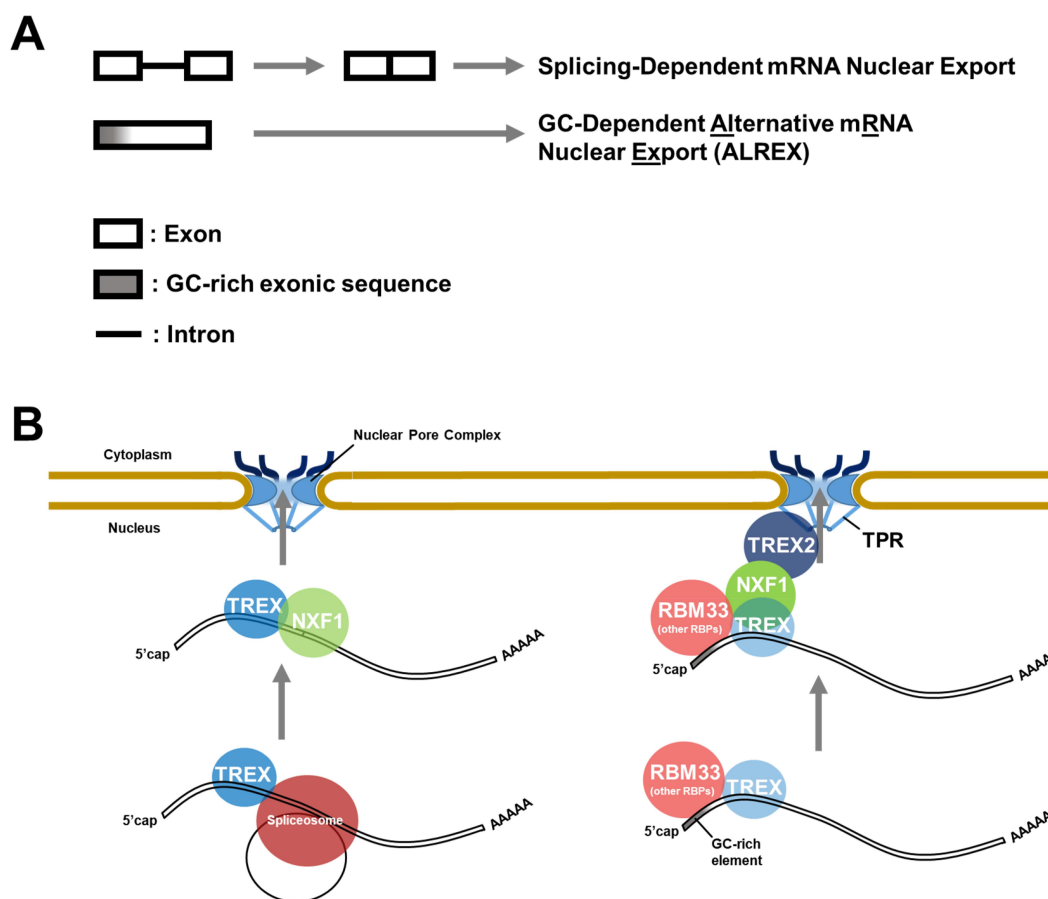


Figure 1. Two mRNA identity features promote the nuclear export of most mRNAs. A) Schematic of the splicing-dependent and GC-dependent mRNA export pathways. B) Illustration of how each feature is recognized by trans-factors, which promote mRNA export. Note that splicing, and likely GC-rich regions, recruit the EJC to both types of mRNA (see text for details).

The nucleocytoplasmic divide is also used in genome defence. It can be used to limit the replication of transposable elements, especially those that have an RNA intermediate. Thus, it acts to reduce the deleteriousness of these selfish bits of DNA [40]. Although this increases the fitness of the organism, the reduction in deleteriousness also suppresses the elimination of transposable elements by purging selection.

Finally, the nucleocytoplasmic divide also plays a role in anti-viral defence in that RNAs that do not have a nuclear history tend to activate innate immune responses [40,41]. In response, many viruses try to translocate to the nucleus in order to replicate. Thus, the nucleocytoplasmic trafficking machinery is monitored by the innate immune response to detect viral invasion. In response, many viruses evolved mechanisms to counteract this by altering aspects of either the nuclear pore or nucleocytoplasmic trafficking [41].

Despite all of these additional features, the primary purpose of the nucleocytoplasmic divide is in mRNA metabolic quality control. In the next sections, we will focus on features that distinguish RNA molecules that have functionally relevant information (mostly mRNAs) from misprocessed and junk RNAs.

Splicing: a key mRNA identity feature that promotes nuclear export

The best-characterized mRNA identity feature that promotes RNA stability and efficient nuclear export is splicing (Figure 1A). It has been widely appreciated that mRNAs from intron-containing genes are more efficiently exported than versions of the exact same mRNA produced from cDNA (i.e. intronless versions of the gene) [21,22]. How splicing and mRNA nuclear export are coupled has been well characterized and many of the molecular details of this process (Figure 1B, Table 1) are understood and reviewed elsewhere [3,5,75,76]. Upon the completion of splicing, the spliceosome helps to recruit the transcription export (TREX) complex and the exon junction complex (EJC) to the mature mRNA [42–45]. TREX in turn helps to promote efficient mRNA export by recruiting the nuclear transport receptor, composed of NXF1 and NXT1

[77–81]. The EJC has been linked to a number of mRNA metabolic steps including an increase in translation, and the removal of misprocessed mRNAs by nonsense-mediated decay [46–48]. Although the exact role of the EJC in mRNA export has remained unclear, it can bind to TREX components, such as ALYREF [46,82,83]. In the past few years, how TREX and the EJC form the core components of the messenger ribonucleoprotein (mRNP) complex has begun to be elucidated by single-particle imaging, X-ray crystallography and Cryo-EM [49,84–86].

Besides its role in export, exon junction density (a measure of how many splicing events give rise to a particular mRNA) is known to be one of the most important contributors to mRNA stability as assessed in a number of transcriptome-wide analyses [87–90]. In particular, deep neural networks were used to show that ORF exon density is strongly associated with increased steady-state mRNA abundance in humans [89]. More recently, a meta-analysis of transcriptome-wide mRNA decay rates (from experiments in 39 human, and 27 mouse, cell lines) also found that ORF exon junction density was the dominant feature for predicting mRNA half-life [90]. Other mRNA processing events, such as 5' capping and polyadenylation, likely also contribute to the deposition of TREX and other nuclear export factors [91–94]; however, unlike splicing, these other processing events in isolation are insufficient to promote efficient nuclear export [5].

GC-Content: a second mRNA identity feature that promotes nuclear export

The second identity feature that promotes nuclear export, is high GC-content at the 5' end of the mRNA (Figure 1A) [27]. Intriguingly, most human protein-coding mRNAs have elevated levels of GC-content at their 5' end regardless of whether they contain introns or not. An illustration of this is shown in Figure 2A where the average GC-content of all human genes containing five exons is plotted. Note that GC-content is highest in the first exon and decreases with every subsequent exon until it dips at the 3' end of the gene.

Table 1. Cis-elements and features that regulate mRNA quality control.

Cis-element or feature	Associated RNAs	Trans-factors	Effect on nuclear export	Other effects
Exon-exon junctions	Spliced protein-coding mRNAs	TREX Complex, EJC [42–45]	Promotes nuclear export [21,22]	Promotes stability, translation and nonsense-mediated decay [46–48]
GC-rich 5' ends (ALREX elements)	Protein-coding mRNAs	RBM33, SARN/THO1/CIP29 (TREX), TPR, SR proteins?, TREX2? [26, 49–51]	Promotes nuclear export [3,23–27]	Promotes stability and translation [25,52,53]
Intact 5'SS motifs	Misprocessed mRNAs (RNAs with retained introns, IPA transcripts) and lncRNAs	ZFC3H1, U1 snRNP [31,54]	Promotes nuclear retention (in conjunction with m6A) [30,54]	Promotes nuclear RNA decay, inhibits 3' cleavage and polyadenylation [30,55–63]
m6A	Poorly spliced mRNAs, long exons, transposable element-derived RNAs	YTH domain-containing proteins [54]	Promotes nuclear retention (in conjunction with 5'SS motifs) [54]	Promotes nuclear and cytoplasmic RNA decay and heterochromatin silencing [64–67]
A-rich sequences	Exons, transposable element-derived RNAs	HUSH Complex	Unknown	Promotes nuclear RNA decay and heterochromatin silencing [68–70]
U-rich sequences	Introns	Unknown	Unknown	Relieves silencing by the HUSH Complex? [68]
dsRNA (A to I editing by ADAR)	Transposable element-derived RNAs, viral RNAs	Major I-binding RBP is unclear, Staufen (dsRNA)	Promotes nuclear retention (A to I editing) [71,72]	Promotes RNA decay (Staufen) [73,74]

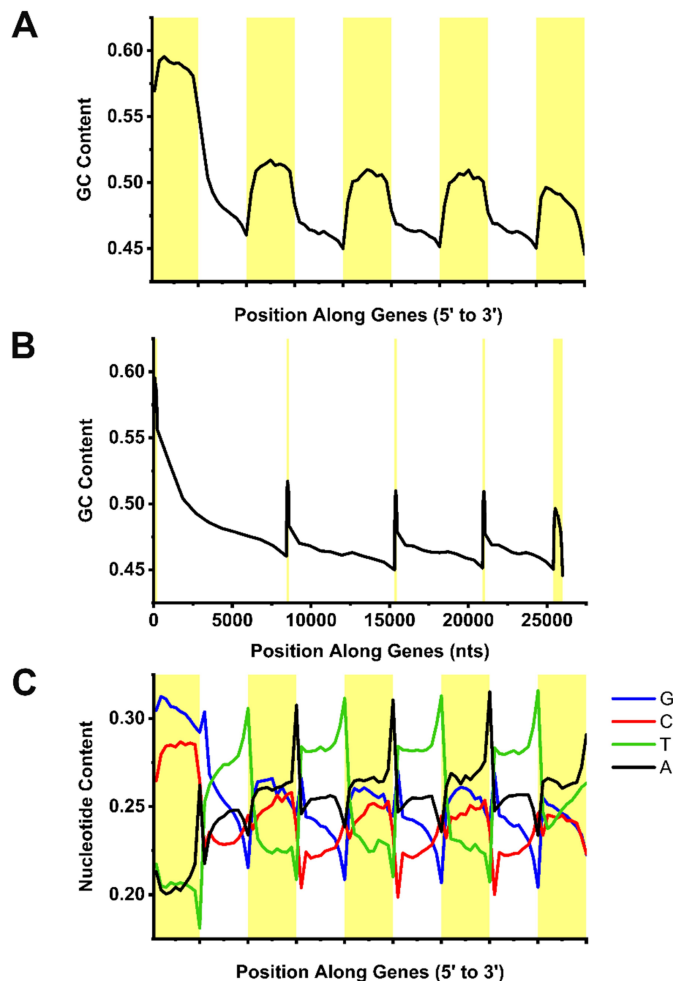


Figure 2. Sequence features of mRNAs. A-B) GC-content averaged over each exon (yellow) and intron (white) of all human protein-coding genes with 5 exons plotted from 5' to 3' ends. Note that in (A) each exon and intron metaplot was normalized, while in (B) they were adjusted to reflect the average length of each exon and intron of all genes in the dataset. C) Similar to (A) except that the average nucleotide-content of the coding strand was plotted from 5' to 3' ends.

Introns, in contrast, are GC-poor, albeit higher than the genomic average, which is 41% GC-content for the human genome. Also, note that the exon and intron sizes in Figure 2A have all been normalized, but in reality, exons are much smaller. Indeed if we replot GC-content but take into account exon and intron sizes (Figure 2B), exons appear as GC-rich islands in a GC-poor sea.

Over 10 years ago, we proposed that mRNAs generated from intronless genes required high GC-content at their 5' end to be efficiently exported from the nucleus to the cytoplasm [3]. In particular, we found that many signal sequence coding regions (SSCRs), which code for the peptide sequences that direct nascent polypeptides to the secretory pathway, also act as RNA elements that promote mRNA nuclear export of intronless reporter mRNAs [23]. These elements had long stretches lacking adenines, tended to be present in the first exons, and contained GC-rich elements [23,24,95]. The depletion of adenines in SSCRs is due to the enrichment between synonymous codons for those that lack adenine and between biochemically similar amino acids for those that have adenine-poor codons

[23,24,95,96]. Importantly, ALREX-elements only promoted export when inserted into the 5' end of reporter mRNAs [97]. This activity was also seen in mitochondrial targeting sequence coding regions (MSCRs) and 'cytoplasmic accumulation RNA' elements (CAR-Es) found in mRNAs from naturally intronless genes [95,98,99]. Again, all of these elements tended to be GC-rich and supported the idea of an alternative RNA export pathway that was sequence-dependent.

More recently, a library of mRNAs, which all coded for the exact same GFP polypeptide, but whose choice of synonymous codons were randomized, were used to identify sequence features that boosted protein expression [25]. It was determined that for intronless mRNAs, high GC-content at the 5' end of the mRNA increased expression, and this was mostly due to increases in the efficiency of mRNA nuclear export. Interestingly, GC-content had little effect in an mRNA library that contained an intron. In parallel, another study identified elements from intronless genes that could promote the export of reporter mRNAs [26]. In agreement with the CAR-E studies, these elements were GC-rich and present near the 5' end of certain intronless mRNAs and lncRNAs, such as *NORAD*. Furthermore, a recent pre-print in BioRxiv used machine learning to show that GC-content was one of the main drivers of high mRNA nuclear export rates [100].

The mechanism of how high GC-content at the 5' end of RNAs promotes export has been elucidated to some extent (Figure 1B, Table 1). It was found that certain components of the TREX complex, in particular the RNA helicase UAP56 (also known as DDX39B) and its paralogue URH49 (also known as DDX39A), were required for the export of reporters with SSCRs or CAR-Es [98,99,101,102]. More recently, it was shown that depletion of the TREX component SARNP (also known as THO1 or CIP29), also had a drastic effect on the export of GC-rich mRNAs [49]. Despite this, other TREX components, such as ALYREF, do not appear to be required for GC-dependent export [51,101]. mRNAs that use the ALREX pathway also appear to require the nuclear pore basket protein TPR and use the nuclear transport receptor NXF1 to cross the nuclear pore [23,26,50,103]. Other RNA-binding factors were identified to associate with CAR-Es; however, whether these recognize GC-rich RNA and promote export remained unclear [99,104]. A recent CRISPR screen for factors that promote the nuclear export of the *NORAD* lncRNA identified RBM33 as being required for the export of intronless GC-rich mRNAs and lncRNAs [51]. This RNA binding protein not only recognizes GC-rich elements but also directly interacts with TREX components, including UAP56, ALYREF, and the nuclear transport receptor NXF1. Another complex, TREX-2, likely acts in the GC-dependent pathway as it functions with TPR [103,105]. Although many nuclear export factors are shared between the splicing and GC-dependent pathway, it appears that the depletion of any given factor tends to have a greater impact on one pathway over the other. For example, depletion of TREX components tends to have greater effects on the splicing-dependent pathway, while depletion of NXF1 and TPR has greater effects on the GC-dependent pathway [26,50,103]. These trends may be

the result of competition by various mRNAs for the remaining export factors when any given component is removed.

Despite all these advances, it remains unclear how at the molecular level GC-rich regions are recognized. Despite the fact that RNA hybridization is energetically favoured in GC-rich RNA, human 5'UTRs, where most of the GC-rich sequence in mRNAs is concentrated, are not enriched for secondary structures in comparison to ORFs and 3'UTRs as assessed by chemical probes [106] and computational modelling [107]. Despite this, there appears to be a slight increase in RNA structures just upstream of the start codon and a depletion in RNA structures at the start codon [106]. Thus, it remains possible that GC-rich regions form structures that are recognized by particular RNA binding proteins. RNA structural features in pre-miRNAs and tRNAs are recognized by their nuclear transport receptors [108,109], and this may be equally true for GC-rich mRNA, but this needs further investigation.

GC-content likely impacts other aspects of mRNA biology. After exon-density, high GC-content in the 5' UTR is the most strongly associated with increased steady-state mRNA abundance in both the human and mouse [89]. There is also some evidence that high GC-content at the 5' end may also enhance an mRNA's translation efficiency [25,52,96], and this may require interactions with RanBP2/Nup358, a component of the cytoplasmic filaments of the nuclear pore complex [52,110]. Interestingly, it had been observed that a subset of EJCs bind to non-canonical sites beyond simply exon-exon junctions [82]. Some of these sites were exclusively present in the first exon of spliced mRNAs and were enriched in the exact same GC-rich motif present in SSCRs [95], suggesting that the EJC may boost the translation of certain GC-rich mRNAs. GC-content may also directly enhance the efficiency of translation elongation. When synonymous human codons are compared, common codons tend to be GC-rich, and thus may be associated with higher rates of translation elongation [111]. GC-rich ORFs also protect mRNAs against cytoplasmic decay [53,111,112]. Despite all these findings, it has been found that selection between codons for translation optimality in humans is weak and the codon distribution in human protein-coding genes is mostly mediated by non-adaptive evolutionary processes, such as GC-biased gene conversion, which elevates local GC-content [113]. Indeed, it is likely that non-adaptive forces, such as GC-biased gene conversion and mutational bias, act in conjunction with adaptive forces to maintain elevated GC-content at the 5' end of most protein coding genes [27,114].

In summary, it has become clear that GC-content is a major determinant of mRNA nuclear export and stability in mammalian cells and that these likely access particular proteins dedicated to this pathway, such as RBM33, and other proteins that are also recruited to spliced mRNAs, such as UAP56, SARNP and NXF1.

Beyond GC-content: other nucleotide-level features of protein-coding genes

Although the GC-content of protein-coding exons has been well documented, there are other features of genes that have received less attention. For example, in human protein-coding genes the nucleotide content differs

substantially between the coding and template strands and this varies along the gene length. This strand asymmetry can be visualized by plotting the individual nucleotide content on the coding strand and can be clearly seen in the collection of human protein-coding genes with five exons (Figure 2C). Some notable trends include a skew towards G and away from C within the first exon (and to a lesser extent in all other exons); a large skew towards A and away from T in internal exons; and extreme strand asymmetry within introns, especially towards T and away from A. This last asymmetry is all the more remarkable when one considers that introns are under a minimal amount of selection and that they are on average at least an order of magnitude longer than exons.

How do these patterns relate to our general understanding of mRNA metabolism? Note that the nucleotide content along the coding strand, which is shown in Figure 2C, matches the nucleotide content along the pre-mRNA. It has been observed that high GC-content in exons and low GC-content in introns may help promote proper splicing or influence how the pre-mRNA is spliced [115–117]. It is possible that other features, such as high T-content in intronic sequences, could be used to identify introns, which have correspondingly high U-content within pre-mRNAs. Indeed, it has been recently observed that U- and A-content are major determinants of how RNA derived from certain transposable elements are identified by the Human Silencing Hub (HUSH) complex (Table 1) [68,69]. It was found that A-rich nascent transcripts that are derived from LINE1 transposable elements can recruit HUSH complex, which in turn targets the RNA for decay and modifies chromatin to enforce genomic silencing. Interestingly, HUSH-mediated RNA decay and silencing can be overridden by inserting an intron into the transcribed region [68]. This likely explains why the HUSH complex does not silence most protein-coding genes, despite the fact that their internal exons are relatively A-rich (Figure 2C). Importantly, the ability of introns to suppress HUSH was not due to the recruitment of the spliceosome as introns that lacked splice site motifs still evaded HUSH silencing [68]. Since intron sequences are T-rich (Figure 2C), and the reverse complement of normally silenced transposable elements (which would also be T-rich) was also found to evade HUSH silencing, it is possible that T- and A-content are nucleotide features that can be used to distinguish protein-coding genes (whose pre-mRNA would be U-rich) from certain transposable elements (which would be transcribed into A-rich RNAs). In line with this, HUSH complex is recruited to mRNAs from intronless genes [68] and to long exons [70], which would represent long tracks of A-rich sequence, a characteristic of most exons in protein coding genes (Figure 2C).

Beyond the biases in T/U- and A-content between introns and exons, it is possible that other subtle biases in the nucleotide content of transcripts may affect RNA metabolism. For example, human genomes are depleted in CpGs due to mutational decay [118] and CpG-rich RNAs appear to be targeted for destruction and this may play a role in anti-viral defense [119].

Nucleotide trends in human genes are not widely appreciated by most molecular biologists, and their effects on RNA metabolism have not been extensively studied. This is fertile ground for future research.

Intact splicing sequences: signals for quality control

Inevitably, splicing errors result in the failure to remove sequences from the transcripts that normally activate splicing. This includes the 5' Splice Site (5'SS), which demarcates the boundary between the exon and an intron, the 3' Splice Site (3'SS), which demarcates the boundary between the intron and the exon, and the branchpoint adenine, which is used to form the lariat structure during splicing. The most typical types of errors are due to splicing failure, the use of cryptic alternative splice sites, and the use of cryptic 3' cleavage/polyadenylation sites (PASs) in introns. In many cases, this results in the presence of intact splicing signals that are typically removed during proper intron removal.

It has been known for quite some time that the presence of intact 5'SS motifs in an mRNA triggers the inhibition of both 3' cleavage and polyadenylation [55–60]. Indeed, when the 5'SS emerges from RNA Polymerase II it directly recruits the U1 snRNP which suppresses the activity of PASs to prevent the premature truncation of newly made transcripts [61–63]. U1 snRNP directly interacts with RNA Polymerase II and this may sterically inhibit the recruitment of the 3' cleavage/polyadenylation machinery [120]. The inhibition of premature cleavage by 5'SS motifs plays a critical role in establishing the 'U1-PAS' axis and preventing the expression of upstream antisense transcripts from

promoters, which have a tendency of activating transcription bidirectionally [63]. However, even in the presence of a 5'SS motif, certain cryptic 3' cleavage/polyadenylation signals may be strong enough to be used at a certain frequency to generate intronic polyadenylated (IPA) transcripts (Figure 3). In these cases, the intact 5'SS in these IPA transcripts triggers their nuclear retention and decay [30].

The molecular mechanism by which intact 5'SS promotes nuclear retention and decay is currently being elucidated (Figure 3, Table 1). This activity requires U1 snRNP and ZFC3H1, a zinc finger containing protein [31]. ZFC3H1 is part of the Poly(A) Exosome Targeting (PAXT) complex which comprises MTR4 (an RNA helicase that targets RNAs to the nuclear exosome), PABPN1 (the nuclear poly(A) binding protein) and several other components [121–124]. This complex has been implicated in the degradation of many types of RNAs by targeting them to the nuclear exosome, the main RNase in the nucleus [121]; however, it appears that certain PAXT components, like ZFC3H1, also act to prevent nuclear export of RNAs that escape degradation [31]. Although MTR4 is not required for the nuclear retention of RNAs with intact 5'SS motifs, it may inhibit the nuclear export of other RNAs by preventing the recruitment of ALYREF, a component of TREX [125]. As for PABPN1, it likely contributes to the nuclear retention of RNAs with intact 5'SS motifs by binding to the poly(A)-tail, however this activity is hard to detect as PABPN1 also promotes RNA nuclear export [94] and these two activities may cancel each other out [31].

The PAXT complex is conserved in eukaryotes, and in fission yeast (*S. pombe*) the equivalent Mtl1-Red1 Core (MTREC) complex is responsible for the nuclear retention and degradation of

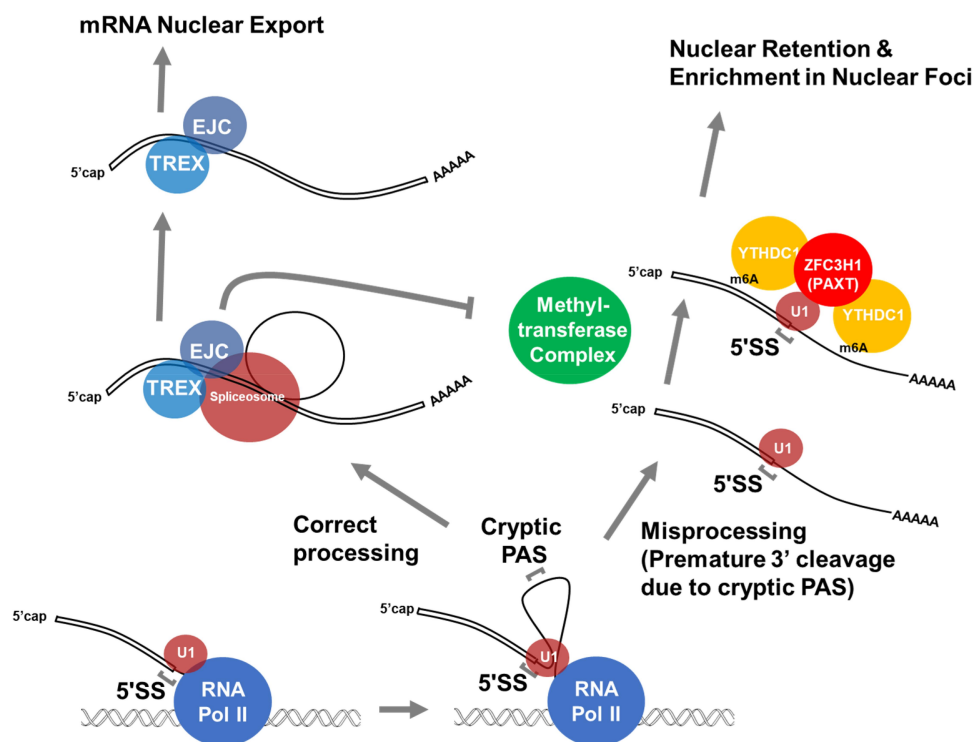


Figure 3. Misprocessing results in the preservation of splicing signals, which promote nuclear retention. Properly processed mRNAs are compared to misprocessed mRNAs that generate IPA transcripts. Note that the IPA transcript contains both an intact 5'SS due to the failure of splicing, and m6A modifications, due to the lack of deposited EJCs, which normally inhibits m6A modifications around the splice site. This could be due to the EJC sterically preventing the methyltransferase from accessing the mRNA (as depicted in the figure) or by the recruitment of demethylases such as ALKBH5.

unspliced mRNAs, transposable element-derived mRNAs, and certain unstable non-coding RNAs [126–128]. Note that Mtl1 and Red1 are the *S. pombe* homologs of MTR4 and ZFC3H1, respectively. In both yeast and humans, ZFC3H1 binds to, and works in concert with, YTH-domain containing proteins to promote the nuclear retention of its substrate transcripts [54,126]. Although human YTH-domain containing proteins (YTHDC1 and YTHDC2) bind to *N*-6-methyladenosine (m6A) RNA modifications, in *S. pombe* the homolog (Mmi1) binds to a specific sequence motif (determinant of selective removal, or DRS) and not m6A [129]. Mmi1 also recruits 3' cleavage machinery to nascent transcripts, and likely activates premature cleavage and decay of certain mRNAs and non-coding RNAs [130].

In both humans and fission yeast, RNAs that are retained by PAXT accumulate in nuclear foci. In *S. pombe*, these foci contain Red1, Mmi1 and RNA substrates [126,128,131]. In humans, RNAs with intact 5'SS motifs are first directed to nuclear speckles [30,31]. Since these structures have been implicated in post-transcriptional splicing [132], this initial targeting may help to complete the removal of introns with weak signals. However, in the case of IPA transcripts, splicing cannot be completed as they lack a branch point and a 3'SS, and these RNAs are transferred to adjacent foci enriched in YTHDC1 [54]. It is likely that these structures also form in cells when there is a general increase in m6A-enriched mRNAs [133], or when PAXT-substrates accumulate due to inhibition of the exosome [124,134]. Both YTH proteins and ZFC3H1 have intrinsically disordered regions that can form biomolecular condensates *in vitro* and may form the matrix of these foci [124,135,136].

Although the nuclear retention of RNAs with 5'SS likely evolved to prevent the export of misprocessed mRNAs, this system was likely co-opted by certain nuclear lncRNAs to ensure their proper localization to the nucleus [137,138]. Unlike mRNAs, lncRNAs are not depleted of 5'SS motifs in their terminal exon [30]. Furthermore, many lncRNAs are poorly spliced [139–142], and the degree to which any lncRNA is nuclear is largely dependent on whether they contain a poorly spliced intron [143], which likely triggers nuclear retention through the presence of intact 5'SS motifs. It has also been observed that lncRNA introns are often spliced using a variety of nearby 5'SS and 3'SS motifs, often leaving behind intact splice signals in the final product [142].

It is important to recognize that many annotated lncRNAs may actually be non-functional transcripts whose deleteriousness is blunted by the fact that they are retained in the nucleus and degraded [6]. Due to their reduced deleteriousness, genomic regions that produce non-functional transcripts are not effectively eliminated by natural selection, and thus we expect to see these accumulate in genomes that are under weak selection regimes, like in most multicellular eukaryotes [4]. This may be enhanced by the proliferation of transposable elements which contain promoter-like sequences that promote the transcription of intergenic regions [12,144]. Indeed, a recent study indicated that the presence of intact 5'SS motifs may help to reduce the deleteriousness of transcripts from genomic loci that eventually evolve into *de novo* lncRNA genes [145]. By triggering the decay of these intermediates and preventing them from being exported, RNA quality control lowers their potential deleteriousness and allows the loci to explore sequence space for extended

periods of time [12]. Despite this, most of the available data suggest that the vast majority of these intermediates eventually lose their ability to be transcribed due to mutational decay, and thus only a vanishing small minority eventually evolve into new non-coding genes.

It is likely that other features of introns, like intact 3'SS motifs, inhibit nuclear mRNA export by recruiting a subset of spliceosome components [146–148]. How these other cis-elements promote nuclear retention will surely be the topic of future investigations.

m6A: another layer of quality control

The *N*-6 methylation of adenine in mRNA has been a topic of intense study, although its exact role in mRNA nuclear export has been unclear due to conflicting findings. m6A is the most prevalent mRNA modification and is nonuniformly distributed across transcripts. It is enriched in the 3' end, surrounding the stop codon, in the 3' UTR and within long internal exons [149,150]. Recent studies have indicated that the m6A modification is excluded around splice sites by the action of the EJC [151–153] (Figure 3). It remains unclear whether the EJC sterically inhibits the m6A methylase complex or promotes m6A removal by recruiting demethylases such as ALKBH5, which has been reported to associate with several components of the EJC [154]. No matter how splicing affects m6A deposition, it is clear that an absence of m6A indicates that a particular transcript is well spliced, and the paucity of this modification, especially within the 5' UTR and ORF, may act as an additional mRNA identity element.

Initial studies indicated that m6A methylation may promote mRNA nuclear export [155–158], however more recent studies have found that this modification tends to repress export and promote RNA decay. In particular, reducing m6A methylation by depleting the methylase (METTL3) elevated the levels of RNAs produced from intergenic regions and transposable elements [64,156,159]. Indeed, the analysis of transcriptomics data by machine learning found that the level of m6A modification of an mRNA correlates with slower nuclear mRNA export rates [100].

In mammals, m6A and m6A-binding proteins are required for the suppression of RNAs generated from transposable elements and unstable ncRNAs [64]. Other studies have shown that m6A promotes mRNA decay in the cytoplasm [65–67]. This may represent a fail-safe mechanism to destroy non-functional RNAs that are poorly spliced, and this may have been co-opted to also target the decay of certain mRNAs. m6A likely has drastic effects on RNA metabolism in the nucleus. It has been linked to the sequestration of *myc* mRNA into nuclear foci [133], and as described in the previous section, we found that the m6A modification was required for the nuclear retention of mis-spliced mRNAs [54] (Figure 3). Other groups have found that YTHDC1 interacts with components of the nuclear exosome targeting (NEXT) complex to target non-coding RNAs for decay [64]. The NEXT and PAXT complexes both share MTR4, although the former acts on RNAs before they are processed while the latter acts on transcripts after they are polyadenylated [121,160]. The NEXT complex also acts with HUSH to silence certain transposable elements [161]. There are likely other connections between m6A and mRNA metabolism. For example, it has been

recently observed that the mRNA nuclear export, ZC3H14 or dNab2, may actively suppress m6A deposition [162], and this could in theory further promote nuclear export.

Overall, the evidence seems quite clear that m6A acts as a layer of quality control to modify poorly processed RNAs, thus marking them as likely non-functional and thus promoting their nuclear retention and destruction.

Adenine to inosine: a quality control mechanism to retain viral RNAs

The last feature that we will discuss is the nuclear retention of double-stranded RNAs (dsRNA, see Table 1). Many RNA viruses must transiently exist as dsRNAs to replicate their genome. These dsRNAs are recognized in the cytoplasm by a host of antiviral sensors that trigger innate immune response pathways [41]. dsRNAs can also be detected in the nucleus by RNA specific adenosine deaminase (ADAR), which converts adenines in the double strand to inosines [163], which in turn promotes nuclear retention [71,72]. As with other nuclear retention pathways, endogenous RNA substrates for ADAR tend to be transcribed from transposable elements [164–168]. In many cases, these endogenous transcripts contain two transposable element-derived sequences that are in reverse orientation from each other, which pair up to form a segment of dsRNA. Other quality control pathways, such as staufen mediated decay (SMD), may also promote the decay of endogenous transcripts that contain dsRNA regions [73,74]. Like other quality control pathways, ADAR-catalysed nuclear retention has been co-opted to regulate the expression of certain mRNAs [169].

As is the case with m6A, inosine-containing RNAs accumulate in nuclear foci. In the case of inosine-containing RNAs, these are paraspeckles [170,171]. Indeed, cells that lack paraspeckles do not have robust nuclear retention of double stranded RNA [171]. The level of ADAR activity may be further regulated by several different RNA binding proteins, and these may also play a role in the regulation of mRNA nuclear retention [172].

Conclusion

Over the past few decades, the field has begun to understand how functionally processed mRNAs are distinguished from mis-processed and spurious transcripts. The field has also dissected how the major molecular machineries that are involved in mRNA nuclear export and retention act to distinguish these two classes of RNAs. There are, however, other features in RNA transcripts which are likely evaluated. We are only beginning to understand what these are and how they are regulated by quality control machineries. Some of these, including motifs [173] and nucleotide modifications [174], have not been confirmed by unbiased whole transcriptome analyses and await independent verification. In addition, it is clear that a major target of nuclear retention and decay are transposable element-derived RNAs, which may be recognized by particular motifs [175]. There are also other less defined nuclear retention elements [29,176,177] which will require further investigation to elucidate how they work at the molecular level.

Acknowledgments

We thank E.S. Lee for comments on the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by a grant to AFP from the Natural Sciences and Engineering Research Council of Canada (FN 492860).

ORCID

Alexander F. Palazzo  <http://orcid.org/0000-0002-9700-1995>

References

- [1] Martin W, Koonin EV. Introns and the origin of nucleus-cytosol compartmentalization. *Nature*. 2006;440(7080):41–45. doi: 10.1038/nature04531
- [2] López-García P, Moreira D. Selective forces for the origin of the eukaryotic nucleus. *BioEssays*. 2006;28(5):525–533. doi: 10.1002/bies.20413
- [3] Palazzo AF, Akef A. Nuclear export as a key arbiter of “mRNA identity” in eukaryotes. *Biochim Biophys Acta*. 2012;1819(6):566–577. doi: 10.1016/j.bbaggm.2011.12.012
- [4] Koonin EV. Splendor and misery of adaptation, or the importance of neutral null for understanding evolution. *BMC Biol*. 2016;14(1):114. doi: 10.1186/s12915-016-0338-2
- [5] Palazzo AF, Lee ES. Sequence determinants for nuclear retention and cytoplasmic export of mRNAs and lncRNAs. *Front Genet*. 2018;9:440. doi: 10.3389/fgene.2018.00440
- [6] Palazzo AF, Lee ES. Non-coding RNA: what is functional and what is junk? *Front Genet*. 2015;6:2. doi: 10.3389/fgene.2015.00002
- [7] Rigau M, Juan D, Valencia A, et al. Intronic CNVs and gene expression variation in human populations. *PLoS Genet*. 2019;15(1):e1007902. doi: 10.1371/journal.pgen.1007902
- [8] Djebali S, Davis CA, Merkel A, et al. Landscape of transcription in human cells. *Nature*. 2012;489(7414):101–108. doi: 10.1038/nature11233
- [9] Palazzo AF, Gregory TR. The case for junk DNA. *PLoS Genet*. 2014;10(5):e1004351. doi: 10.1371/journal.pgen.1004351
- [10] Ponting C. Biological function in the twilight zone of sequence conservation. *BMC Biol*. 2017;15(1):71. doi: 10.1186/s12915-017-0411-5
- [11] Ponting CP, Haerty W. Genome-wide analysis of human long noncoding RNAs: a provocative review. *Ann Rev Genomics Hum Genet*. 2022;23(1):153–172. doi: 10.1146/annurev-genom-112921-123710
- [12] Palazzo AF, Koonin EV. Functional long non-coding RNAs evolve from junk transcripts. *Cell*. 2020;183(5):1–12. doi: 10.1016/j.cell.2020.09.047
- [13] Lee H, Zhang Z, Krause HM. Long noncoding RNAs and repetitive elements: junk or intimate evolutionary partners? *Trends Genet*. 2019;35(12):892–902. doi: 10.1016/j.tig.2019.09.006
- [14] Mattick JS, Dinger ME. The extent of functionality in the human genome. *Hugo J*. 2013;7(1):2. doi: 10.1186/1877-6566-7-2
- [15] López-García P, Moreira D. Open questions on the origin of eukaryotes. *Trends Ecol Evol*. 2015;30(11):697–708. doi: 10.1016/j.tree.2015.09.005
- [16] Sharp PA. Five easy pieces. *Science*. 1991;254(5032):663. doi: 10.1126/science.1948046
- [17] Smathers CM, Robart AR. The mechanism of splicing as told by group II introns: ancestors of the spliceosome. *Biochim Biophys*

- Acta Gene Regul Mech. 2019;1862(11):194390. doi: [10.1016/j.bbagr.2019.06.001](https://doi.org/10.1016/j.bbagr.2019.06.001)
- [18] Vosseberg J, Stolker D, von der Dunk SHA, et al. Integrating phylogenetics with intron positions illuminates the origin of the complex spliceosome. *Mol Biol Evol*. 2023;40(1):msad011. doi: [10.1093/molbev/msad011](https://doi.org/10.1093/molbev/msad011)
- [19] Rogozin IB, Carmel L, Csuros M, et al. Origin and evolution of spliceosomal introns. *Biol Direct*. 2012;7(1):11. doi: [10.1186/1745-6150-7-11](https://doi.org/10.1186/1745-6150-7-11)
- [20] Field MC, Rout MP. Coatomeer in the universe of cellular complexity. *Mol Biol Cell*. 2022;33(14):e8. doi: [10.1091/mbc.E19-01-0012](https://doi.org/10.1091/mbc.E19-01-0012)
- [21] Luo MJ, Reed R. Splicing is required for rapid and efficient mRNA export in metazoans. *Proc Natl Acad Sci, USA*. 1999;96(26):14937–14942. doi: [10.1073/pnas.96.26.14937](https://doi.org/10.1073/pnas.96.26.14937)
- [22] Valencia P, Dias AP, Reed R. Splicing promotes rapid and efficient mRNA export in mammalian cells. *Proc Natl Acad Sci, USA*. 2008;105(9):3386–3391. doi: [10.1073/pnas.0800250105](https://doi.org/10.1073/pnas.0800250105)
- [23] Palazzo AF, Springer M, Shibata Y, et al. The signal sequence coding region promotes nuclear export of mRNA. *PLoS Biol*. 2007;5(12):e322. doi: [10.1371/journal.pbio.0050322](https://doi.org/10.1371/journal.pbio.0050322)
- [24] Palazzo A, Mahadevan K, Tarnawsky S. ALREX-elements and introns: two identity elements that promote mRNA nuclear export. *WIREs RNA*. 2013;4(5):523–533. doi: [10.1002/wrna.1176](https://doi.org/10.1002/wrna.1176)
- [25] Mordstein C, Savisaar R, Young RS, et al. Codon usage and splicing jointly influence mRNA localization. *Cell Systems*. 2020;10(4):351–362.e8. doi: [10.1016/j.cels.2020.03.001](https://doi.org/10.1016/j.cels.2020.03.001)
- [26] Zuckerman B, Ron M, Mikl M, et al. Gene architecture and sequence composition underpin selective dependency of nuclear export of long RNAs on NXF1 and the TREX complex. *Molecular Cell*. 2020;79(2):251–267.e6. doi: [10.1016/j.molcel.2020.05.013](https://doi.org/10.1016/j.molcel.2020.05.013)
- [27] Palazzo AF, Kang YM. GC-content biases in protein-coding genes act as an “mRNA identity” feature for nuclear export. *BioEssays*. 2021;43(2):2000197. doi: [10.1002/bies.202000197](https://doi.org/10.1002/bies.202000197)
- [28] Garland W, Jensen TH. Nuclear sorting of RNA. *WIREs RNA*. 2020;11(2):e1572. doi: [10.1002/wrna.1572](https://doi.org/10.1002/wrna.1572)
- [29] Akef A, Lee ES, Palazzo AF. Splicing promotes the nuclear export of β -globin mRNA by overcoming nuclear retention elements. *RNA*. 2015;21(11):1908–1920. doi: [10.1261/rna.051987.115](https://doi.org/10.1261/rna.051987.115)
- [30] Lee ES, Akef A, Mahadevan K, et al. The consensus 5' splice site motif inhibits mRNA nuclear export. *PLoS One*. 2015;10(3):e0122743. doi: [10.1371/journal.pone.0122743](https://doi.org/10.1371/journal.pone.0122743)
- [31] Lee ES, Smith HW, Wolf EJ, et al. ZFC3H1 and U1-70K promote the nuclear retention of mRNAs with 5' splice site motifs within nuclear speckles. *RNA*. 2022;28(6):878–894. doi: [10.1261/rna.079104.122](https://doi.org/10.1261/rna.079104.122)
- [32] Culjkovic B, Topisirovic I, Skrabanek L, et al. eIF4E promotes nuclear export of cyclin D1 mRNAs via an element in the 3'UTR. *J Cell Bio*. 2005;169(2):245–256. doi: [10.1083/jcb.200501019](https://doi.org/10.1083/jcb.200501019)
- [33] Yamazaki T, Fujiwara N, Yukinaga H, et al. The closely related RNA helicases, UAP56 and URH49, preferentially form distinct mRNA export machineries and coordinately regulate mitotic progression. *Mol Biol Cell*. 2010;21(16):2953–2965. doi: [10.1091/mbc.E09-10-0913](https://doi.org/10.1091/mbc.E09-10-0913)
- [34] Borden K, Culkovic-Kraljic B. mRNA export and its dysregulation in disease. In: Yang W, editor. *Nuclear-cytoplasmic transport. Nucleic acids and molecular biology*. Springer International Publishing; 2018. p. 179–204. doi: [10.1007/978-3-319-77309-4_8](https://doi.org/10.1007/978-3-319-77309-4_8)
- [35] Lefauieux D, Sen S, Jiang K, et al. Kinetics of mRNA nuclear export regulate innate immune response gene expression. *Nat Commun*. 2022;13(1):7197. doi: [10.1038/s41467-022-34635-5](https://doi.org/10.1038/s41467-022-34635-5)
- [36] Zander G, Hackmann A, Bender L, et al. mRNA quality control is bypassed for immediate export of stress-responsive transcripts. *Nature*. 2016;540(7634):593–596. doi: [10.1038/nature20572](https://doi.org/10.1038/nature20572)
- [37] Mihaylov SR, Castelli LM, Lin YH, et al. The master energy homeostasis regulator PGC-1 α exhibits an mRNA nuclear export function. *Nat Commun*. 2023;14(1):5496. doi: [10.1038/s41467-023-41304-8](https://doi.org/10.1038/s41467-023-41304-8)
- [38] Boutz PL, Bhutkar A, Sharp PA. Detained introns are a novel, widespread class of post-transcriptionally spliced introns. *Genes Dev*. 2015;29(1):63–80. doi: [10.1101/gad.247361.114](https://doi.org/10.1101/gad.247361.114)
- [39] Bahar Halpern K, Caspi I, Lemze D, et al. Nuclear retention of mRNA in mammalian tissues. *Cell Rep*. 2015;13(12):2653–2662. doi: [10.1016/j.celrep.2015.11.036](https://doi.org/10.1016/j.celrep.2015.11.036)
- [40] Madhani HD. The frustrated gene: origins of eukaryotic gene expression. *Cell*. 2013;155(4):744–749. doi: [10.1016/j.cell.2013.10.003](https://doi.org/10.1016/j.cell.2013.10.003)
- [41] Shen Q, Wang YE, Palazzo AF. Crosstalk between nucleocytoplasmic trafficking and the innate immune response to viral infection. *J Biol Chem*. 2021;297(1):100856. doi: [10.1016/j.jbc.2021.100856](https://doi.org/10.1016/j.jbc.2021.100856)
- [42] Luo ML, Zhou Z, Magni K, et al. Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature*. 2001;413(6856):644–647. doi: [10.1038/35098106](https://doi.org/10.1038/35098106)
- [43] Strässer K, Masuda S, Mason P, et al. TREX is a conserved complex coupling transcription with messenger RNA export. *Nature*. 2002;417(6886):304–308. doi: [10.1038/nature746](https://doi.org/10.1038/nature746)
- [44] Masuda S, Das R, Cheng H, et al. Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev*. 2005;19(13):1512–1517. doi: [10.1101/gad.1302205](https://doi.org/10.1101/gad.1302205)
- [45] Le Hir H, Izaurralde E, Maquat LE, et al. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J*. 2000;19(24):6860–6869. doi: [10.1093/emboj/19.24.6860](https://doi.org/10.1093/emboj/19.24.6860)
- [46] Le Hir H, Gatfield D, Izaurralde E, et al. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J*. 2001;20(17):4987–4997. doi: [10.1093/emboj/20.17.4987](https://doi.org/10.1093/emboj/20.17.4987)
- [47] Nott A, Le Hir H, Moore MJ. Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. *Genes Dev*. 2004;18(2):210–222. doi: [10.1101/gad.1163204](https://doi.org/10.1101/gad.1163204)
- [48] Woodward LA, Mabin JW, Gangras P, et al. The exon junction complex: a lifelong guardian of mRNA fate. *WIREs RNA*. 2017;8(3). doi: [10.1002/wrna.1411](https://doi.org/10.1002/wrna.1411)
- [49] Xie Y, Gao S, Zhang K, et al. Structural basis for high-order complex of SARNP and DDX39B to facilitate mRNP assembly. *Cell Rep*. 2023;42(8):112988. doi: [10.1016/j.celrep.2023.112988](https://doi.org/10.1016/j.celrep.2023.112988)
- [50] Lee ES, Wolf EJ, Ihn SSJ, et al. TPR is required for the efficient nuclear export of mRNAs and lncRNAs from short and intron-poor genes. *Nucleic Acids Res*. 2020;48(20):11645–11663. doi: [10.1093/nar/gkaa919](https://doi.org/10.1093/nar/gkaa919)
- [51] Thomas A, Rehfeld F, Zhang H, et al. RBM33 directs the nuclear export of transcripts containing GC-rich elements. *Genes Dev*. 2022;36(9–10):550–565. doi: [10.1101/gad.349456.122](https://doi.org/10.1101/gad.349456.122)
- [52] Mahadevan K, Zhang H, Akef A, et al. RanBP2/Nup358 potentiates the translation of a subset of mRNAs encoding secretory proteins. *PLoS Biol*. 2013;11(4):e1001545. doi: [10.1371/journal.pbio.1001545](https://doi.org/10.1371/journal.pbio.1001545)
- [53] Kudla G, Lipinski L, Caffin F, et al. High guanine and cytosine content increases mRNA levels in mammalian cells. *PLoS Biol*. 2006;4(6):e180. doi: [10.1371/journal.pbio.0040180](https://doi.org/10.1371/journal.pbio.0040180)
- [54] Lee ES, de Olivera LS, Jomphe RY, et al. N-6-methyladenosine (m6A) promotes the nuclear retention of mRNAs with intact 5' splice site motifs. Published online 2023 Jun 21; 2023.06.20.545713. doi: [10.1101/2023.06.20.545713](https://doi.org/10.1101/2023.06.20.545713)
- [55] Furth PA, Choe WT, Rex JH, et al. Sequences homologous to 5' splice sites are required for the inhibitory activity of papillomavirus late 3' untranslated regions. *Mol Cell Biol*. 1994;14(8):5278–5289. doi: [10.1128/MCB.14.8.5278](https://doi.org/10.1128/MCB.14.8.5278)
- [56] Ashe MP, Griffin P, James W, et al. Poly(a) site selection in the HIV-1 provirus: inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. *Genes Dev*. 1995;9(23):3008–3025. doi: [10.1101/gad.9.23.3008](https://doi.org/10.1101/gad.9.23.3008)
- [57] Ashe MP, Pearson LH, Proudfoot NJ. The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream

- major splice donor site. *EMBO J.* **1997**;16(18):5752–5763. doi: [10.1093/emboj/16.18.5752](https://doi.org/10.1093/emboj/16.18.5752)
- [58] Gunderson SI, Polycarpou-Schwarz M, Mattaj IW. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol Cell.* **1998**;1(2):255–264. doi: [10.1016/S1097-2765\(00\)80026-X](https://doi.org/10.1016/S1097-2765(00)80026-X)
- [59] Vagner S, Rügsegger U, Gunderson SI, et al. Position-dependent inhibition of the cleavage step of pre-mRNA 3'-end processing by U1 snRNP. *RNA.* **2000**;6(2):178–188. doi: [10.1017/S1355838200991854](https://doi.org/10.1017/S1355838200991854)
- [60] Langemeier J, Schrom EM, Rabner A, et al. A complex immunodeficiency is based on U1 snRNP-mediated poly(A) site suppression. *EMBO J.* **2012**;31(20):4035–4044. doi: [10.1038/emboj.2012.252](https://doi.org/10.1038/emboj.2012.252)
- [61] Kaida D, Berg MG, Younis I, et al. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature.* **2010**;468(7324):664–668. doi: [10.1038/nature09479](https://doi.org/10.1038/nature09479)
- [62] Berg MG, Singh LN, Younis I, et al. U1 snRNP determines mRNA length and regulates isoform expression. *Cell.* **2012**;150(1):53–64. doi: [10.1016/j.cell.2012.05.029](https://doi.org/10.1016/j.cell.2012.05.029)
- [63] Almada AE, Wu X, Kriz AJ, et al. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature.* **2013**;499(7458):360–363. doi: [10.1038/nature12349](https://doi.org/10.1038/nature12349)
- [64] Liu J, Dou X, Chen C, et al. N6-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science.* **2020**;367(6477):580–586. doi: [10.1126/science.aay6018](https://doi.org/10.1126/science.aay6018)
- [65] Wang X, Lu Z, Gomez A, et al. m6A-dependent regulation of messenger RNA stability. *Nature.* **2014**;505(7481):117–120. doi: [10.1038/nature12730](https://doi.org/10.1038/nature12730)
- [66] Du H, Zhao Y, He J, et al. YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. *Nat Commun.* **2016**;7(1):12626. doi: [10.1038/ncomms12626](https://doi.org/10.1038/ncomms12626)
- [67] Boo SH, Ha H, Lee Y, et al. UPF1 promotes rapid degradation of m6A-containing RNAs. *Cell Rep.* **2022**;39(8):110861. doi: [10.1016/j.celrep.2022.110861](https://doi.org/10.1016/j.celrep.2022.110861)
- [68] Seczynska M, Bloor S, Cuesta SM, et al. Genome surveillance by HUSH-mediated silencing of intronless mobile elements. *Nature.* **2022**;601(7893):440–445. doi: [10.1038/s41586-021-04228-1](https://doi.org/10.1038/s41586-021-04228-1)
- [69] Seczynska M, Lehner PJ. The sound of silence: mechanisms and implications of HUSH complex function. *Trends Genet.* **2023**;39(4):251–267. doi: [10.1016/j.tig.2022.12.005](https://doi.org/10.1016/j.tig.2022.12.005)
- [70] Spencley AL, Bar S, Swigut T, et al. Co-transcriptional genome surveillance by HUSH is coupled to termination machinery. *Molecular Cell.* **2023**;83(10):1623–1639.e8. doi: [10.1016/j.molcel.2023.04.014](https://doi.org/10.1016/j.molcel.2023.04.014)
- [71] Kumar M, Carmichael GG. Nuclear antisense RNA induces extensive adenosine modifications and nuclear retention of target transcripts. *PNAS.* **1997**;94(8):3542–3547. doi: [10.1073/pnas.94.8.3542](https://doi.org/10.1073/pnas.94.8.3542)
- [72] Zhang Z, Carmichael GG. The fate of dsRNA in the nucleus. *Cell.* **2001**;106(4):465–476. doi: [10.1016/S0092-8674\(01\)00466-4](https://doi.org/10.1016/S0092-8674(01)00466-4)
- [73] Park E, Maquat LE. Staufen-mediated mRNA decay. *WIREs RNA.* **2013**;4(4):423–435. doi: [10.1002/wrna.1168](https://doi.org/10.1002/wrna.1168)
- [74] Lucas BA, Lavi E, Shiue L, et al. Evidence for convergent evolution of SINE-directed Staufen-mediated mRNA decay. *PNAS.* **2018**;115(5):968–973. doi: [10.1073/pnas.1715531115](https://doi.org/10.1073/pnas.1715531115)
- [75] Katahira J. mRNA export and the TREX complex. *Biochim Biophys Acta Gene Regul Mech.* **2012**;1819(6):507–513. doi: [10.1016/j.bbagr.2011.12.001](https://doi.org/10.1016/j.bbagr.2011.12.001)
- [76] Heath CG, Viphakone N, Wilson SA. The role of TREX in gene expression and disease. *Biochem J.* **2016**;473(19):2911–2935. doi: [10.1042/BCJ20160010](https://doi.org/10.1042/BCJ20160010)
- [77] Stutz F, Bachi A, Doerks T, et al. REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA.* **2000**;6(4):638–650. doi: [10.1017/S1355838200000078](https://doi.org/10.1017/S1355838200000078)
- [78] Hautbergue GM, Hung ML, Golovanov AP, et al. Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. *Proc Natl Acad Sci, USA.* **2008**;105(13):5154–5159. doi: [10.1073/pnas.0709167105](https://doi.org/10.1073/pnas.0709167105)
- [79] Katahira J, Inoue H, Hurt E, et al. Adaptor aly and co-adaptor Thoc5 function in the Tap-p15-mediated nuclear export of HSP70 mRNA. *EMBO J.* **2009**;28(5):556–567. doi: [10.1038/emboj.2009.5](https://doi.org/10.1038/emboj.2009.5)
- [80] Hung ML, Hautbergue GM, Snijders APL, et al. Arginine methylation of REF/ALY promotes efficient handover of mRNA to TAP/NXF1. *Nucleic Acids Res.* **2010**;38(10):3351–3361. doi: [10.1093/nar/gkq033](https://doi.org/10.1093/nar/gkq033)
- [81] Viphakone N, Hautbergue GM, Walsh M, et al. TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nat Commun.* **2012**;3(1):1006. doi: [10.1038/ncomms2005](https://doi.org/10.1038/ncomms2005)
- [82] Singh G, Kucukural A, Cenik C, et al. The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell.* **2012**;151(4):915–916. doi: [10.1016/j.cell.2012.10.032](https://doi.org/10.1016/j.cell.2012.10.032)
- [83] Viphakone N, Sudbery I, Griffith L, et al. Co-transcriptional loading of RNA export factors shapes the human transcriptome. *Molecular Cell.* **2019**;75(2):310–323.e8. doi: [10.1016/j.molcel.2019.04.034](https://doi.org/10.1016/j.molcel.2019.04.034)
- [84] Xie Y, Clarke BP, Kim YJ, et al. Cryo-EM structure of the yeast TREX complex and coordination with the SR-like protein Gbp2. *Elife.* **2021**;10:e65699. doi: [10.7554/eLife.65699](https://doi.org/10.7554/eLife.65699)
- [85] Pacheco-Fiallos B, Vorländer MK, Riabov-Bassat D, et al. mRNA recognition and packaging by the human transcription–export complex. *Nature.* **2023**;616(7958):828–835. doi: [10.1038/s41586-023-05904-0](https://doi.org/10.1038/s41586-023-05904-0)
- [86] Asada R, Dominguez A, Montpetit B. Single-molecule quantitation of RNA-binding protein occupancy and stoichiometry defines a role for Yra1 (Aly/REF) in nuclear mRNP organization. *Cell Rep.* **2023**;42(11):113415. doi: [10.1016/j.celrep.2023.113415](https://doi.org/10.1016/j.celrep.2023.113415)
- [87] Sharova LV, Sharov AA, Nedorezov T, et al. Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. *DNA Res.* **2009**;16(1):45–58. doi: [10.1093/dnares/dsn030](https://doi.org/10.1093/dnares/dsn030)
- [88] Spies N, Burge CB, Bartel DP. 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. *Genome Res.* **2013**;23(12):2078–2090. doi: [10.1101/gr.156919.113](https://doi.org/10.1101/gr.156919.113)
- [89] Agarwal V, Shendure J. Predicting mRNA abundance directly from genomic sequence using deep convolutional neural networks. *Cell Rep.* **2020**;31(7):107663. doi: [10.1016/j.celrep.2020.107663](https://doi.org/10.1016/j.celrep.2020.107663)
- [90] Agarwal V, Kelley D. The genetic and biochemical determinants of mRNA degradation rates in mammals. *Genome Biology.* **2022**;23(2):245. doi: [10.1186/s13059-022-02811-x](https://doi.org/10.1186/s13059-022-02811-x)
- [91] Green DM, Marfatia KA, Crafton EB, et al. Nab2p is required for poly(A) RNA export in *Saccharomyces cerevisiae* and is regulated by arginine methylation via Hmt1p. *J Biol Chem.* **2002**;277(10):7752–7760. doi: [10.1074/jbc.M110053200](https://doi.org/10.1074/jbc.M110053200)
- [92] Hector RE, Nykamp KR, Dheur S, et al. Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. *EMBO J.* **2002**;21(7):1800–1810. doi: [10.1093/emboj/21.7.1800](https://doi.org/10.1093/emboj/21.7.1800)
- [93] Cheng H, Dufu K, Lee CS, et al. Human mRNA export machinery recruited to the 5' end of mRNA. *Cell.* **2006**;127(7):1389–1400. doi: [10.1016/j.cell.2006.10.044](https://doi.org/10.1016/j.cell.2006.10.044)
- [94] Apponi LH, Leung SW, Williams KR, et al. Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis. *Hum Mol Genet.* **2010**;19(6):1058–1065. doi: [10.1093/hmg/ddp569](https://doi.org/10.1093/hmg/ddp569)
- [95] Cenik C, Chua HN, Zhang H, et al. Genome analysis reveals interplay between 5'UTR introns and nuclear mRNA export for secretory and mitochondrial genes. *PLoS Genet.* **2011**;7(4):e1001366. doi: [10.1371/journal.pgen.1001366](https://doi.org/10.1371/journal.pgen.1001366)
- [96] Cenik C, Chua HN, Singh G, et al. A common class of transcripts with 5'-intron depletion, distinct early coding sequence features, and N(1)-methyladenosine modification. *RNA.* **2017**;23(3):270–283. doi: [10.1261/rna.059105.116](https://doi.org/10.1261/rna.059105.116)
- [97] Tarnawsky SP, Palazzo AF. Positional requirements for the stimulation of mRNA nuclear export by ALREX-promoting elements. *Mol Biosyst.* **2012**;8(10):2527–2530. doi: [10.1039/c2mb25016k](https://doi.org/10.1039/c2mb25016k)
- [98] Lei H, Dias AP, Reed R. Export and stability of naturally intronless mRNAs require specific coding region sequences and the

- TREX mRNA export complex. *Proc Natl Acad Sci, USA.* 2011;108(44):17985–17990. doi: [10.1073/pnas.1113076108](https://doi.org/10.1073/pnas.1113076108)
- [99] Lei H, Zhai B, Yin S, et al. Evidence that a consensus element found in naturally intronless mRNAs promotes mRNA export. *Nucleic Acids Res.* 2013;41(4):2517–2525. doi: [10.1093/nar/gks1314](https://doi.org/10.1093/nar/gks1314)
- [100] Smalec BM, Ietswaart R, Choquet K, et al. Genome-wide quantification of RNA flow across subcellular compartments reveals determinants of the mammalian transcript life cycle. Published online 2022 Aug 21;2022.08.21.504696. doi: [10.1101/2022.08.21.504696](https://doi.org/10.1101/2022.08.21.504696)
- [101] Akef A, Zhang H, Masuda S, et al. Trafficking of mRNAs containing ALREX-promoting elements through nuclear speckles. *Nucleus.* 2013;4(4):326–340. doi: [10.4161/nuc.26052](https://doi.org/10.4161/nuc.26052)
- [102] Chi B, Wang K, Du Y, et al. A sub-element in PRE enhances nuclear export of intronless mRNAs by recruiting the TREX complex via ZC3H18. *Nucleic Acids Res.* 2014;42(11):7305–7318. doi: [10.1093/nar/gku350](https://doi.org/10.1093/nar/gku350)
- [103] Aksenova V, Smith A, Lee H, et al. Nucleoporin TPR is an integral component of the TREX-2 mRNA export pathway. *Nat Commun.* 2020;11(1):4577. doi: [10.1038/s41467-020-18266-2](https://doi.org/10.1038/s41467-020-18266-2)
- [104] Folco EG, Lee CS, Dufu K, et al. The proteins PDIP3 and ZC11A associate with the human TREX complex in an ATP-dependent manner and function in mRNA export. *PLoS One.* 2012;7(8):e43804. doi: [10.1371/journal.pone.0043804](https://doi.org/10.1371/journal.pone.0043804)
- [105] Wickramasinghe VO, Andrews R, Ellis P, et al. Selective nuclear export of specific classes of mRNA from mammalian nuclei is promoted by GANP. *Nucleic Acids Res.* 2014;42(8):5059–5071. doi: [10.1093/nar/gku095](https://doi.org/10.1093/nar/gku095)
- [106] Wan Y, Qu K, Zhang QC, et al. Landscape and variation of RNA secondary structure across the human transcriptome. *Nature.* 2014;505(7485):706–709. doi: [10.1038/nature12946](https://doi.org/10.1038/nature12946)
- [107] Kawaguchi R, Kiryu H. Parallel computation of genome-scale RNA secondary structure to detect structural constraints on human genome. *BMC Bioinf.* 2016;17(1):203. doi: [10.1186/s12859-016-1067-9](https://doi.org/10.1186/s12859-016-1067-9)
- [108] Cook AG, Fukuhara N, Jinek M, et al. Structures of the tRNA export factor in the nuclear and cytosolic states. *Nature.* 2009;461(7260):60–65. doi: [10.1038/nature08394](https://doi.org/10.1038/nature08394)
- [109] Okada C, Yamashita E, Lee SJ, et al. A high-resolution structure of the pre-microRNA nuclear export machinery. *Science.* 2009;326(5957):1275–1279. doi: [10.1126/science.1178705](https://doi.org/10.1126/science.1178705)
- [110] Bley CJ, Nie S, Mobbs GW, et al. Architecture of the cytoplasmic face of the nuclear pore. *Science.* 2022;376(6598):eabm9129. doi: [10.1126/science.abm9129](https://doi.org/10.1126/science.abm9129)
- [111] Courel M, Clément Y, Bossevain C, et al. GC content shapes mRNA storage and decay in human cells. *Elife.* 2019;8:e49708. doi: [10.7554/eLife.49708](https://doi.org/10.7554/eLife.49708)
- [112] Jowhar Z, Xu A, Venkataraman S, et al. A ubiquitous GC content signature underlies multimodal mRNA regulation by DDX3X. Published online 2023 Nov 23; 2023.05.11.540322. doi: [10.1101/2023.05.11.540322](https://doi.org/10.1101/2023.05.11.540322)
- [113] Pouyet F, Mouchiroud D, Duret L, et al. Recombination, meiotic expression and human codon usage. *Elife.* 2017;6:e27344. doi: [10.7554/eLife.27344](https://doi.org/10.7554/eLife.27344)
- [114] Palazzo AF, Kejiou NS. Non-Darwinian molecular biology. *Front Genet.* 2022;13:831068. doi: [10.3389/fgene.2022.831068](https://doi.org/10.3389/fgene.2022.831068)
- [115] Amit M, Donyo M, Hollander D, et al. Differential GC content between exons and introns establishes distinct strategies of splice-site recognition. *Cell Rep.* 2012;1(5):543–556. doi: [10.1016/j.celrep.2012.03.013](https://doi.org/10.1016/j.celrep.2012.03.013)
- [116] Tammer L, Hameiri O, Keydar I, et al. Gene architecture directs splicing outcome in separate nuclear spatial regions. *Molecular Cell.* 2022;82(5):1021–1034.e8. doi: [10.1016/j.molcel.2022.02.001](https://doi.org/10.1016/j.molcel.2022.02.001)
- [117] Barutcu AR, Wu M, Braunschweig U, et al. Systematic mapping of nuclear domain-associated transcripts reveals speckles and lamina as hubs of functionally distinct retained introns. *Molecular Cell.* 2022;82(5):1035–1052.e9. doi: [10.1016/j.molcel.2021.12.010](https://doi.org/10.1016/j.molcel.2021.12.010)
- [118] Bird AP. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.* 1980;8(7):1499–1504. doi: [10.1093/nar/8.7.1499](https://doi.org/10.1093/nar/8.7.1499)
- [119] Takata MA, Gonçalves-Carneiro D, Zang TM, et al. CG dinucleotide suppression enables antiviral defence targeting non-self RNA. *Nature.* 2017;550(7674):124–127. doi: [10.1038/nature24039](https://doi.org/10.1038/nature24039)
- [120] Zhang S, Aibara S, Vos SM, et al. Structure of a transcribing RNA polymerase II–U1 snRNP complex. *Science.* 2021;371(6526):305–309. doi: [10.1126/science.abf1870](https://doi.org/10.1126/science.abf1870)
- [121] Meola N, Domanski M, Karadoulama E, et al. Identification of a nuclear exosome decay pathway for processed transcripts. *Mol Cell.* 2016;64(3):520–533. doi: [10.1016/j.molcel.2016.09.025](https://doi.org/10.1016/j.molcel.2016.09.025)
- [122] Ogami K, Richard P, Chen Y, et al. An Mtr4/ZFC3H1 complex facilitates turnover of unstable nuclear RNAs to prevent their cytoplasmic transport and global translational repression. *Genes Dev.* 2017;31(12):1257–1271. doi: [10.1101/gad.302604.117](https://doi.org/10.1101/gad.302604.117)
- [123] Silla T, Schmid M, Dou Y, et al. The human ZC3H3 and RBM26/27 proteins are critical for PAXT-mediated nuclear RNA decay. *Nucleic Acids Res.* 2020;48(5):2518–2530. doi: [10.1093/nar/gkz1238](https://doi.org/10.1093/nar/gkz1238)
- [124] Wang Y, Fan J, Wang J, et al. ZFC3H1 prevents RNA trafficking into nuclear speckles through condensation. *Nucleic Acids Res.* 2021;49(18):10630–10643. doi: [10.1093/nar/gkab774](https://doi.org/10.1093/nar/gkab774)
- [125] Fan J, Kuai B, Wu G, et al. Exosome cofactor hMTR4 competes with export adaptor ALYREF to ensure balanced nuclear RNA pools for degradation and export. *EMBO J.* 2017;36(19):2870–2886. doi: [10.15252/embj.201696139](https://doi.org/10.15252/embj.201696139)
- [126] Sugiyama T, Sugioka-Sugiyama R. Red1 promotes the elimination of meiosis-specific mRNAs in vegetatively growing fission yeast. *EMBO J.* 2011;30(6):1027–1039. doi: [10.1038/emboj.2011.32](https://doi.org/10.1038/emboj.2011.32)
- [127] Zhou Y, Zhu J, Schermann G, et al. The fission yeast MTREC complex targets CUTs and unspliced pre-mRNAs to the nuclear exosome. *Nat Commun.* 2015;6(1):7050. doi: [10.1038/ncomms8050](https://doi.org/10.1038/ncomms8050)
- [128] Shichino Y, Otsubo Y, Yamamoto M, et al. Meiotic gene silencing complex MTREC/NURS recruits the nuclear exosome to YTH-RNA-binding protein Mmi1. *PLoS Genet.* 2020;16(2):e1008598. doi: [10.1371/journal.pgen.1008598](https://doi.org/10.1371/journal.pgen.1008598)
- [129] Wang C, Zhu Y, Bao H, et al. A novel RNA-binding mode of the YTH domain reveals the mechanism for recognition of determinant of selective removal by Mmi1. *Nucleic Acids Res.* 2016;44(2):969–982. doi: [10.1093/nar/gkv1382](https://doi.org/10.1093/nar/gkv1382)
- [130] Vo TV, Dhakshnamoorthy J, Larkin M, et al. CPF recruitment to non-canonical transcription termination sites triggers heterochromatin assembly and gene silencing. *Cell Rep.* 2019;28(1):267–281. doi: [10.1016/j.celrep.2019.05.107](https://doi.org/10.1016/j.celrep.2019.05.107)
- [131] Harigaya Y, Tanaka H, Yamanaka S, et al. Selective elimination of messenger RNA prevents an incidence of untimely meiosis. *Nature.* 2006;442(7098):45–50. doi: [10.1038/nature04881](https://doi.org/10.1038/nature04881)
- [132] Dias AP, Dufu K, Lei H, et al. A role for TREX components in the release of spliced mRNA from nuclear speckle domains. *Nat Commun.* 2010;1(1):97. doi: [10.1038/ncomms1103](https://doi.org/10.1038/ncomms1103)
- [133] Cheng Y, Xie W, Pickering BF, et al. N6-methyladenosine on mRNA facilitates a phase-separated nuclear body that suppresses myeloid leukemic differentiation. *Cancer Cell.* 2021;39(7):958–972.e8. doi: [10.1016/j.ccell.2021.04.017](https://doi.org/10.1016/j.ccell.2021.04.017)
- [134] Silla T, Karadoulama E, Măkosa D, et al. The RNA exosome adaptor ZFC3H1 functionally competes with nuclear export activity to retain target transcripts. *Cell Rep.* 2018;23(7):2199–2210. doi: [10.1016/j.celrep.2018.04.061](https://doi.org/10.1016/j.celrep.2018.04.061)
- [135] Gao Y, Pei G, Li D, et al. Multivalent m6A motifs promote phase separation of YTHDF proteins. *Cell Res.* 2019;29(9):767–769. doi: [10.1038/s41422-019-0210-3](https://doi.org/10.1038/s41422-019-0210-3)
- [136] Wang J, Wang L, Diao J, et al. Binding to m6A RNA promotes YTHDF2-mediated phase separation. *Protein Cell.* 2020;11(4):304–307. doi: [10.1007/s13238-019-00660-2](https://doi.org/10.1007/s13238-019-00660-2)
- [137] Azam S, Hou S, Zhu B, et al. Nuclear retention element recruits U1 snRNP components to restrain spliced lncRNAs in the nucleus. *RNA Biol.* 2019;16(8):1001–1009. doi: [10.1080/15476286.2019.1620061](https://doi.org/10.1080/15476286.2019.1620061)
- [138] Yin Y, Lu JY, Zhang X, et al. U1 snRNP regulates chromatin retention of noncoding RNAs. *Nature.* 2020;580(7801):147–150. doi: [10.1038/s41586-020-2105-3](https://doi.org/10.1038/s41586-020-2105-3)

- [139] Tilgner H, Knowles DG, Johnson R, et al. Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res.* 2012;22(9):1616–1625. doi: [10.1101/gr.134445.111](#)
- [140] Melé M, Mattioli K, Mallard W, et al. Chromatin environment, transcriptional regulation, and splicing distinguish lincRNAs and mRNAs. *Genome Res.* 2017;27(1):27–37. doi: [10.1101/gr.214205.116](#)
- [141] Mukherjee N, Calviello L, Hirsekorn A, et al. Integrative classification of human coding and noncoding genes through RNA metabolism profiles. *Nat Struct Mol Biol.* 2017;24(1):86–96. doi: [10.1038/nsmb.3325](#)
- [142] Deveson IW, Brunck ME, Blackburn J, et al. Universal alternative splicing of noncoding exons. *Cell Syst.* 2018;6(2):245–255.e5. doi: [10.1016/j.cels.2017.12.005](#)
- [143] Zuckerman B, Ulitsky I. Predictive models of subcellular localization of long RNAs. *RNA.* 2019;25(5):557–572. doi: [10.1261/rna.068288.118](#)
- [144] Johnson R, Guigó R. The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. *RNA.* 2014;20(7):959–976. doi: [10.1261/rna.044560.114](#)
- [145] An NA, Zhang J, Mo F, et al. De novo genes with an lncRNA origin encode unique human brain developmental functionality. *Nat Ecol Evol.* 2023;7(2):264–278. doi: [10.1038/s41559-022-01925-6](#)
- [146] Legrain P, Rosbash M. Some cis- and trans-acting mutants for splicing target pre-mRNA to the cytoplasm. *Cell.* 1989;57(4):573–583. doi: [10.1016/0092-8674\(89\)90127-X](#)
- [147] Rain JC, Legrain P. In vivo commitment to splicing in yeast involves the nucleotide upstream from the branch site conserved sequence and the Mud2 protein. *EMBO J.* 1997;16(7):1759–1771. doi: [10.1093/emboj/16.7.1759](#)
- [148] Takemura R, Takeiwa T, Taniguchi I, et al. Multiple factors in the early splicing complex are involved in the nuclear retention of pre-mRNAs in mammalian cells. *Genes Cells.* 2011;16(10):1035–1049. doi: [10.1111/j.1365-2443.2011.01548.x](#)
- [149] Dominissini D, Moshitch-Moshkovitz S, Schwartz S, et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature.* 2012;485(7397):201–206. doi: [10.1038/nature11112](#)
- [150] Molinie B, Giallourakis CC. Genome-wide location analyses of N6-methyladenosine modifications (m6A-seq). *Methods Mol Biol.* 2017;1562:45–53. doi: [10.1007/978-1-4939-6807-7_4](#)
- [151] Yang X, Triboulet R, Liu Q, et al. Exon junction complex shapes the m6A epitranscriptome. *Nat Commun.* 2022;13(1):7904. doi: [10.1038/s41467-022-35643-1](#)
- [152] Uzonyi A, Dierks D, Nir R, et al. Exclusion of m6A from splice-site proximal regions by the exon junction complex dictates m6A topologies and mRNA stability. *Mol Cell.* 2023;83(2):237–251.e7. doi: [10.1016/j.molcel.2022.12.026](#)
- [153] He PC, Wei J, Dou X, et al. Exon architecture controls mRNA m6A suppression and gene expression. *Science.* 2023;379(6633):677–682. doi: [10.1126/science.abj9090](#)
- [154] Covelo-Molares H, Obrdlik A, Poštulková I, et al. The comprehensive interactomes of human adenosine RNA methyltransferases and demethylases reveal distinct functional and regulatory features. *Nucleic Acids Res.* 2021;49(19):10895–10910. doi: [10.1093/nar/gkab900](#)
- [155] Zheng G, Dahl JA, Niu Y, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell.* 2013;49(1):18–29. doi: [10.1016/j.molcel.2012.10.015](#)
- [156] Roundtree IA, Luo GZ, Zhang Z, et al. YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *Elife.* 2017;6:e31311. doi: [10.7554/eLife.31311](#)
- [157] Lesbirel S, Viphakone N, Parker M, et al. The m6A-methylase complex recruits TREX and regulates mRNA export. *Sci Rep.* 2018;8(1):13827. doi: [10.1038/s41598-018-32310-8](#)
- [158] Lesbirel S, Wilson SA. The m6A-methylase complex and mRNA export. *Biochim Biophys Acta Gene Regul Mech.* 2019;1862(3):319–328. doi: [10.1016/j.bbagr.2018.09.008](#)
- [159] He C, Lan F. RNA m6A meets transposable elements and chromatin. *Protein Cell.* 2021;12(12):906–910. doi: [10.1007/s13238-021-00859-2](#)
- [160] Wu G, Schmid M, Rib L, et al. A two-layered targeting mechanism underlies nuclear RNA sorting by the human exosome. *Cell Rep.* 2020;30(7):2387–2401.e5. doi: [10.1016/j.celrep.2020.01.068](#)
- [161] Garland W, Müller I, Wu M, et al. Chromatin modifier HUSH co-operates with RNA decay factor NEXT to restrict transposable element expression. *Molecular Cell.* 2022;82(9):1691–1707.e8. doi: [10.1016/j.molcel.2022.03.004](#)
- [162] Jalloh B, Lancaster CL, Rounds JC, et al. The drosophila Nab2 RNA binding protein inhibits m6A methylation and male-specific splicing of sex lethal transcript in female neuronal tissue. *Elife.* 2023;12:e64904. doi: [10.7554/eLife.64904](#)
- [163] Polson AG, Crain PF, Pomerantz SC, et al. The mechanism of adenosine to inosine conversion by the double-stranded RNA unwinding/modifying activity: a high-performance liquid chromatography-mass spectrometry analysis. *Biochemistry.* 1991;30(49):11507–11514. doi: [10.1021/bi00113a004](#)
- [164] Athanasiadis A, Rich A, Maas S, et al. Widespread A-to-I RNA editing of alu-containing mRNAs in the human transcriptome. *PLoS Biol.* 2004;2(12):e391. doi: [10.1371/journal.pbio.0020391](#)
- [165] Blow M, Futreal PA, Wooster R, et al. A survey of RNA editing in human brain. *Genome Res.* 2004;14(12):2379–2387. doi: [10.1101/gr.2951204](#)
- [166] Kim DDY, Kim TTY, Walsh T, et al. Widespread RNA editing of embedded alu elements in the human transcriptome. *Genome Res.* 2004;14(9):1719–1725. doi: [10.1101/gr.2855504](#)
- [167] Levanon EY, Eisenberg E, Yelin R, et al. Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat Biotechnol.* 2004;22(8):1001–1005. doi: [10.1038/nbt996](#)
- [168] Chen LL, Carmichael GG. Gene regulation by SINES and inosines: biological consequences of A-to-I editing of Alu element inverted repeats. *Cell Cycle.* 2008;7(21):3294–3301. doi: [10.4161/cc.7.21.6927](#)
- [169] Prasanth KV, Prasanth SG, Xuan Z, et al. Regulating gene expression through RNA nuclear retention. *Cell.* 2005;123(2):249–263. doi: [10.1016/j.cell.2005.08.033](#)
- [170] Chen LL, DeCervo JN, Carmichael GG. Alu element-mediated gene silencing. *EMBO J.* 2008;27(12):1694–1705. doi: [10.1038/emboj.2008.94](#)
- [171] Chen LL, Carmichael GG. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Molecular Cell.* 2009;35(4):467–478. doi: [10.1016/j.molcel.2009.06.027](#)
- [172] Freund EC, Sapiro AL, Li Q, et al. Unbiased identification of trans regulators of ADAR and A-to-I RNA editing. *Cell Rep.* 2020;31(7):107656. doi: [10.1016/j.celrep.2020.107656](#)
- [173] Zhang B, Gunawardane L, Niaz F, et al. A novel RNA motif mediates the strict nuclear localization of a long noncoding RNA. *Mol Cell Biol.* 2014;34(12):2318–2329. doi: [10.1128/MCB.01673-13](#)
- [174] Yang X, Yang Y, Sun BF, et al. 5-methylcytosine promotes mRNA export — NSUN2 as the methyltransferase and ALYREF as an m5C reader. *Cell Res.* 2017;27(5):606–625. doi: [10.1038/cr.2017.55](#)
- [175] Lubelsky Y, Ulitsky I. Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature.* 2018;555(7694):107–111. doi: [10.1038/nature25757](#)
- [176] Miyagawa R, Tano K, Mizuno R, et al. Identification of cis- and trans-acting factors involved in the localization of MALAT-1 noncoding RNA to nuclear speckles. *RNA.* 2012;18(4):738–751. doi: [10.1261/rna.028639.111](#)
- [177] Shukla CJ, McCorkindale AL, Gerhardinger C, et al. High-throughput identification of RNA nuclear enrichment sequences. *EMBO J.* 2018;37(6). doi: [10.15252/embj.201798452](#)