How cells get the message: dynamic assembly and function of mRNA–protein complexes

Michaela Müller-McNicoll and Karla M. Neugebauer

Abstract | mRNA is packaged into ribonucleoprotein particles called mRNPs. A multitude of RNA-binding proteins as well as a host of associated proteins participate in the fate of mRNA from transcription and processing in the nucleus to translation and decay in the cytoplasm. Methodological innovations in cell biology and genome-wide high-throughput approaches have revealed an unexpected diversity of mRNA-associated proteins and unforeseen interconnections between mRNA-processing steps. Recent insights into mRNP formation *in vivo* have also highlighted the importance of mRNP packaging, which can sort RNAs on the basis of their length and determine mRNA fate through alternative mRNP assembly, processing and export pathways.

Messenger ribonucleoprotein particles (mRNPs). Complexes composed of mature mRNAs bound by various RNA-binding proteins and associated proteins recruited via protein–protein interactions. The formation of mRNPs allows proper packaging of the mRNA, which is essential for efficient nuclear export.

5′ end capping
As soon as a nascent transcript emerges from RNA polymerase II during transcription, it is capped at its 5′ end. This 7-methylguanosine cap (m^7G cap) protects the mRNA from degradation and is essential for its translation.

Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstrasse 108, 01307 Dresden, Germany.
e-mails: mmueller@mpi-cbg.de; neugebau@mpi-cbg.de
doi:10.1038/nrg3434
Published online 12 March 2013

Messenger RNA (mRNA) is never alone. It is coated and compacted by RNA-binding proteins (RBPs), forming large messenger ribonucleoprotein particles (mRNPs). RBPs assemble on nascent and mature mRNAs. They serve as structural elements for mRNP packaging and modify the output of gene expression at all steps of the mRNA life cycle: transcription, 5′ end capping, precursor mRNA (pre-mRNA) splicing, 3′ end processing, nuclear export, localization, translation and mRNA stability (FIG. 1).

Analysis of the protein composition of RNPs containing small non-coding RNAs (for example, small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs) or microRNAs (miRNAs)) has been a tractable problem[1,2], whereas studying the composition of mRNPs has been more challenging. mRNAs are not only longer, but they also undergo more processing steps that can be differentially regulated among tissues and cellular compartments. A combination of biochemical and cell biological research, recently complemented by genome-wide studies, has led to the current notion that each mRNA is bound by multiple RBPs and that individual RBPs have hundreds to thousands of mRNA targets[3–5]. Furthermore, RBPs are often components of multi-protein complexes and bring additional proteins into mRNPs through protein–protein interactions. It is generally accepted that during their journey from transcription in the nucleus to translation in the cytoplasm, mRNPs continuously change their composition. However, prolonged binding of particular RBPs can couple succeeding steps in gene expression either by directly influencing more than one step of mRNA processing or by affecting downstream events through the recruitment of additional proteins[6]. Interestingly, evidence is accumulating to show that RBPs can also couple non-sequential steps in mRNA metabolism. This suggests that besides continuous mRNP remodelling, alternative ways for mRNP assembly may exist.

In this Review, we focus on recent insights from *in vivo* studies carried out in mammalian cells (except where noted) that have contributed to addressing several major open questions. How many different proteins associate with mRNA in general and/or with individual mRNAs? How important is mRNA length and mRNA packaging for gene expression? To what extent does the protein composition of mRNPs undergo remodelling during the lifetime of an mRNA? How are nuclear and cytoplasmic mRNP maturation steps interconnected by shared RBPs? Owing to space limitations, we do not discuss the coupling of cytoplasmic RNA metabolic steps; the reader is referred to excellent recent reviews on this topic[6,7]. Complementing the topics discussed here, several interesting reviews have discussed RNA-binding specificities and *in vivo* binding motifs of RBPs[5], methods for identifying RNA–protein interactions globally[4,8], mRNP surveillance[9] and nuclear export dynamics of mRNPs[10].

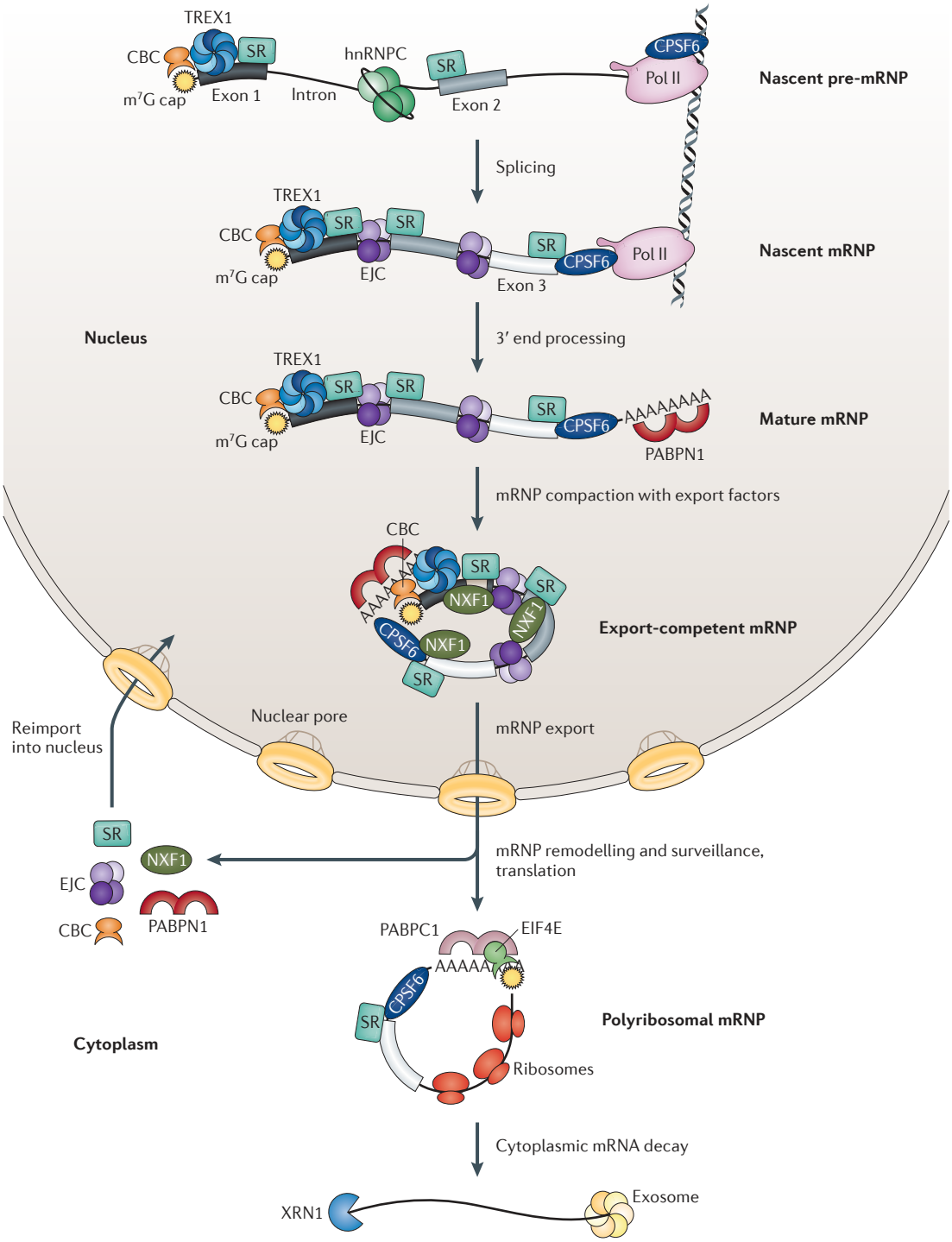

Figure 1 | **mRNPs: co-transcriptional assembly and remodelling from transcription to translation and degradation.** This figure illustrates the current view of co-transcriptional assembly and remodelling of messenger ribonucleoprotein particles (mRNPs) in mammalian cells at various steps throughout their lifetime; some example RNA-binding proteins (RBPs) that are mentioned in the main text are shown. One example is the cap-binding complex (CBC), which binds to the 7-methylguanosine (m^7G) cap and facilitates splicing, 3′ end formation and then export as it diffuses with bound mRNAs to nuclear pores[123]. In the cytoplasm, the CBC is exchanged for the translation initiation factor eukaryotic initiation factor 4E (EIF4E)[7]. Most RBPs that were associated with the mRNA during processing and export are replaced by ribosomes on actively translated mRNAs. CPSF6, cleavage and polyadenylation specificity factor 6; EJC, exon junction complex; hnRNPC, heterogeneous nuclear ribonucleoprotein C; NXF1, nuclear export factor 1; PABPC1, cytoplasmic poly(A) binding protein 1; PABPN1, nuclear poly(A)-binding protein 1; Pol II, RNA polymerase II; pre-mRNP, precursor mRNP; TREX1,transcription export complex 1; XRN1, 5′–3′ exoribonuclease 1.

Peptidyl-prolyl-isomerases (PPIs). A group of metabolic enzymes that catalyse the *cis–trans* isomerization of peptide bonds in polypeptide chains. PPIs have important roles in the folding of newly synthesized proteins but were recently shown also to bind mRNAs.

RNA-binding proteins

The RNA-binding protein repertoire: how many different proteins associate with mRNAs? More than 600 RBPs are annotated in mammalian genomes on the basis of the presence of known RNA-binding domains (RBDs), but the functions and *in vivo* binding specificities of most RBPs are not well understood. Many RBPs have been inferred from sequence homology, and until recently only a small fraction had been validated for RNA binding *in vivo*. An important step towards a comprehensive atlas of proteins that bind to mRNA *in vivo* was the recent establishment of 'interactome capture' and protein-profiling methods[11,12] (BOX 1).

Combining two complementary ultraviolet (UV) crosslinking approaches, one study identified most of the expressed mRNA-binding proteins in HeLa cells. In another study, photoactivatable-ribonucleoside-enhanced UV crosslinking and immunoprecipitation (PAR-CLIP) was used to identify the mRNA-bound proteome in human embryonic kidney (HEK) cells and to map globally the sites of protein–mRNA interactions (BOX 1). The PAR-CLIP study supported the notion that transcripts are generally bound and regulated by a multitude of RBPs[3,12]. Both studies identified similar numbers of RBPs (860 and 797, respectively), which is substantially more than the ~600 known and presumed RBPs. Among these, 315 proteins[11] and 245 proteins[12] were not previously annotated as RNA binding, and ~200 proteins had only been inferred as RBPs by homology. It was further shown that several metabolic enzymes, such as peptidyl-prolyl-isomerases (PPIs), bind RNA *in vivo*[11]. These numbers constitute the most up-to-date estimate of the total number of RBPs that are expressed in a human cell line at a given time, but presumably many more tissue-specific RBPs await discovery. For comparison, 120 mRNA-binding proteins have recently been identified in yeast using a similar technique[13]. Moreover, both of the interactome capture studies investigated only mature mRNPs, hence RBPs bound to pre-mRNA, non-polyadenylated RNA or introns remain elusive. It will be exciting to study the binding specificities of the newly identified RBPs,

Box 1 | Methods for characterizing mRNP composition and functions

Early methods

Diverse biochemical *in vitro* approaches, such as electrophoretic mobility shift assays (EMSA) or ultraviolet (UV) crosslinking of proteins to their target RNAs have been used to study RNA–protein interactions and messenger ribonucleoprotein particle (mRNP) assembly over the past three decades. Twenty years ago, the systematic evolution of ligands by exponential enrichment (SELEX) method was introduced to isolate high-affinity RNA aptamers from highly diverse pools of *in vitro* transcribed RNAs[108]. This method enabled high-affinity RNA recognition elements (RREs) to be identified for numerous RBPs *in vitro*. However, high-affinity recognition might be unfavourable *in vivo*[5].

The yeast three-hybrid system, which was developed in 1996, enabled RNA–protein interactions to be monitored *in vivo* for the first time by measuring the expression levels of reporter genes[109]. Diverse RNA immunoprecipitation (RIP) protocols established in the past 10 years finally permitted the identification of endogenous RNAs present in complex mRNPs[110]. RIP followed by microarray-based identification of protein-bound RNAs (RIP–chip) revealed, for example, that the targets of an RBP could be functionally related transcripts[41,111]. As a reciprocal approach, individual mRNPs have been purified, and their protein composition has been analysed by mass spectrometry[40].

High-throughput methods

More recently, different high-throughput methods based on UV crosslinking, immunoprecipitation and deep sequencing (CLIP–seq) were introduced as powerful methods to determine *in vivo* protein–RNA interactions on a global scale[20,112]. UV crosslinking requires direct contact between protein and RNA and does not promote protein–protein crosslinks. UV light with a wavelength of 254 nm crosslinks the naturally photoreactive nucleotide bases to specific amino acids, and the photoreactive thionucleosides 4-thiouridine (4-SU) and 6-thioguanosine (6-SG) that are used in photoactivatable-ribonucleoside-enhanced CLIP (PAR-CLIP) can be crosslinked at 365 nm. PAR-CLIP and another CLIP variant, termed iCLIP, provide an almost nucleotide resolution of RNA-binding sites[20,113]. The crosslinked amino acid covalently bound to recovered RNA constitutes a barrier for the reverse transcriptase enzyme and either causes specific nucleotide changes (PAR-CLIP) or truncation (iCLIP) of the cDNA during reverse transcription. Some RBPs are favoured by one method, and neither protocol seems to be superior[11,114]. The different CLIP–seq approaches have been successfully used to generate transcriptome-wide RNA maps for numerous RBPs, confirming that one RBP can have few or thousands of mRNA targets. One important question in RNA biology — how many RNAs one particular RBP can bind — can be answered with these methods[4].

An alternative crosslinking protocol termed CLASH (for crosslinking, ligation and sequencing of hybrids) has further revealed tertiary interactions between guide small nucleolar RNAs (snoRNAs) bound to snoRNA-binding proteins and their ribosomal RNA (rRNA) targets[115]. This protocol could be applied to study interactions between Argonaute (AGO) proteins, microRNAs and their target transcripts. The recently developed interactome capture and protein occupancy profiling approaches also involve *in vivo* UV crosslinking followed by capture of polyadenylated RNA bound to proteins on oligo(dT)-coated beads, subsequent stringent washes to eliminate all non-crosslinked proteins and release by nuclease digestion. Released proteins are finally analysed by mass spectrometry[11,12].

Unfortunately, the interactome capture approaches cannot provide the protein composition of distinct mRNPs, which are expressed at a given time or within cellular compartments; instead, all expressed mRNAs and their bound RBPs are simultaneously analysed. Also, some proteins may stay associated with the selected RNA through high-affinity protein–protein interactions that are stable despite high-stringency washing rather than direct RNA–protein binding. Validation — for example, by iCLIP — will continue to be crucial to the interpretation of RNA interactomes.

their endogenous targets and their functions and to identify comprehensively proteins that are bound to non-polyadenylated RNAs.

Novel RNA-binding domains. RBDs are evolutionarily conserved protein domains that recognize specific sequence or structural elements in their target RNAs. More than 40 annotated RBDs exist in the databases (Interpro[14]), but only some of the most frequently occurring, such as the RNA recognition motif (RRM), have been characterized in detail. They often bind to short, degenerate 4–6-nucleotide (nt) segments with a weak affinity in a sequence- and/or structure-specific manner[4,5]. While searching for common motifs among the newly identified RBPs in the interactome capture studies, several non-canonical RBDs were discovered[11,12]. Most interestingly, some of the RBPs contained globular and/or disordered regions enriched in short repetitive amino acid motifs with unusual RBDs[11]. One example is the K patch, which is an unstructured lysine-rich region. This novel RBD occurs in the abundant endoplasmic-reticulum-membrane-bound protein p180, which was recently shown to act as an mRNA receptor that promotes the association of mRNA with the endoplasmic reticulum surface[15]. The K patch in p180 was confirmed to bind RNA directly *in vitro* and *in vivo,* probably through non-specific interactions with the RNA backbone and was required for the efficient anchoring of certain transcripts to the endoplasmic reticulum membrane[11,15]. This example highlights an important function for a novel RBP that contains disordered regions with short repetitive amino acid motifs that are capable of binding RNA.

Multimerization of RNA-binding proteins or multi-domain recognition. Structural biology has provided some molecular details and principles of RNA recognition for single RBDs bound to model transcripts. However, many RBPs require multiple RBDs for their *in vivo* functions, and the interplay between these domains is poorly understood[4,5,16]. RBPs with multiple RBDs can recognize longer stretches of RNA, can combine multiple weak interactions to achieve high affinity and specificity or can create new or extended binding surfaces for RNAs[16]. Multimerization also allows RBPs to recognize RNA sequences that belong to different RNAs or that are separated by long intervening stretches of nucleotides (for example, as occurs in the spliceosome)[17]. Additionally, RNA recognition domains can be coupled with catalytic domains[16]. For example, the multi-domain ribonuclease Dicer from *Giardia intestinalis* functions as a 'molecular ruler' by binding to double-stranded RNA (dsRNA) through its PAZ domain and positioning its catalytic RNase III domain ~25 nt away[18]. The coordinated activity of these two domains accounts for the observed length of small interfering RNAs (siRNAs), which are generated through cleavage of dsRNA by Dicer[18].

Other examples of multi-domain RNA-binding proteins that wrap or loop RNA *in vivo* include polypyrimidine-tract-binding protein (PTB) and insulin-like growth factor 2 mRNA-binding protein 1 (IGF2BP1)[16,19]. Structural analyses of such multi-domain RBPs in combination with high-resolution binding approaches such as iCLIP and PAR-CLIP (BOX 1) now allow the elucidation of combinatorial binding of multiple RBDs and provide insights into the packaging of pre-mRNAs, which were previously elusive[5,20] (discussed further below).

Roles of RBPs in packaging

mRNPs: the importance of RNA length and mRNA packaging for gene expression. Classic biochemical studies revealed that mature mRNAs are packaged into RNA–protein complexes known as mRNPs, whereas nuclear pre-mRNA is packaged into heterogeneous nuclear ribonucleoprotein particles (hnRNPs)[21,22]. Analyses of the average composition and size of purified human hnRNPs and mRNPs have shown that particle size correlates with the length of the RNA component[22,23]. Consistent with these earlier observations, a recent live-cell imaging study using individual, fluorescently labelled mRNPs assembled on RNAs ranging from 1.7 to 17 kb in length showed that nuclear and cytoplasmic diffusion coefficients, which reflect particle size, decrease with RNA length[24]. Interestingly, although the RNA lengths varied tenfold, the range of diffusion coefficients varied only fourfold to fivefold, suggesting that mRNP packaging affects the dynamics of the RNP particle and that longer RNAs are more tightly packaged[24]. This might reflect a cellular strategy used to ensure that longer mRNAs are not disadvantaged regarding their export and subsequent translation.

The hnRNP proteins have a major role in the packaging of RNA in the nucleus. Approximately 20 hnRNP proteins have been identified in purified human hnRNP particles, and they are designated A1 to U[25]. All hnRNP proteins are nuclear, are mostly associated with pre-mRNAs and only minimally bind to mature mRNAs[22,26]. Early studies proposed that hnRNP proteins scaffold pre-mRNA during transcription and act as chaperones and/or regulators during subsequent nuclear events in mRNA biogenesis analogously to the role of histones in DNA packaging[27]. Recent *in vivo* studies, focusing on hnRNPC, have now revisited this proposal.

Nuclear hnRNPC is a multimeric RBP that forms a 3:1 heterotetramer of hnRNPC1 and hnRNPC2 subunits, and each subunit contains one RRM. Two of the four RRMs in the tetramer bind to neighbouring uridine tracts separated by short intervening stretches[28], and 150–250 nt of RNA are wrapped around each hnRNPC tetramer[29,30]. A recent study using *in vivo* crosslinking and immunoprecipitation followed by deep sequencing (iCLIP; BOX 1) revealed that hnRNPC tetramers predominantly bind to the introns of pre-mRNAs (from 55% of all annotated protein-coding genes), suggesting co-transcriptional assembly[20]. The same study confirmed the short spacing between neighbouring hnRNPC binding sites and also found crosslinked uridines with a defined spacing of 165 nt, a finding that is consistent with the length of RNA that can be wrapped by one tetramer. Interestingly, hnRNPC binding occurred in a repeated fashion along the pre-mRNA — particularly

RNA recognition motif
(RRM). One of the most abundant RNA-binding domains in eukaryotes.

Spliceosome
A ribonucleoprotein complex that is responsible for splicing nuclear precursor mRNA (pre-mRNA). It is composed of five small nuclear ribonucleoproteins (snRNPs) and more than 50 non-snRNP proteins, which recognize and assemble on exon–intron boundaries to catalyse intron removal from the precursor mRNA (pre-mRNA).

Dicer
An RNase III family endonuclease that processes double-stranded RNA and precursor microRNAs into small interfering RNAs and microRNAs, respectively.

Heterogeneous nuclear ribonucleoprotein particles
(hnRNPs). Complexes of newly synthesized precursor mRNA (pre-mRNA) and RNA-binding proteins, known as heterogeneous ribonucleoproteins, which form during transcription in the cell nucleus. The abundant hnRNP proteins regulate splicing and mark the RNA as immature. After splicing has occurred, the hnRNP proteins mainly remain bound to spliced introns.

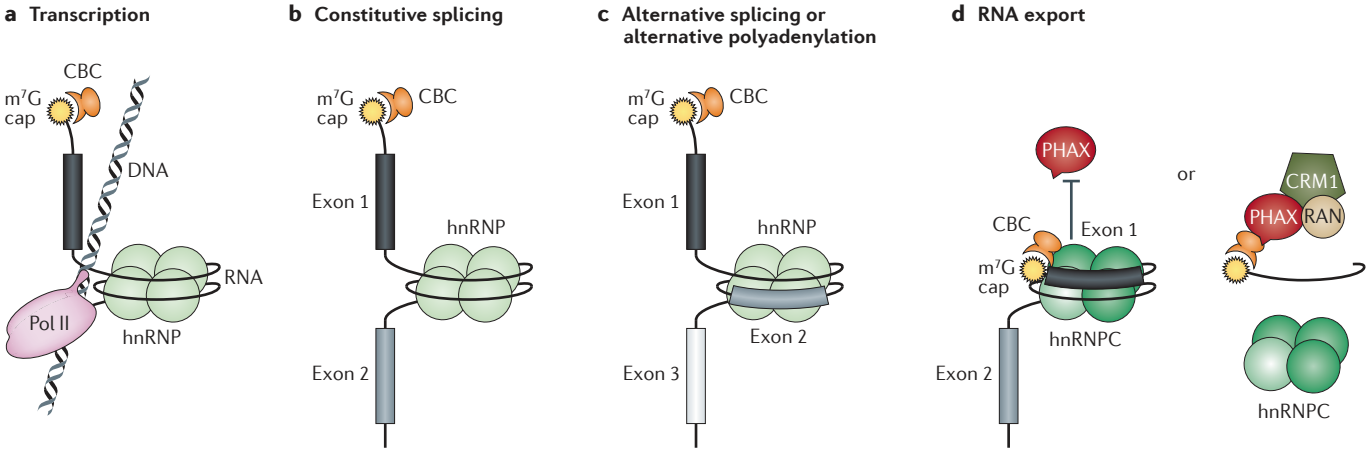

Figure 2 | **Importance of co-transcriptional mRNA packaging during RNA-processing steps from transcription to export.** **a** | Transcription: packaging of nascent RNA prevents physical entanglement with DNA during transcription — for example, the formation of deleterious R loops. **b** | Constitutive splicing: packaging of partially spliced RNA holds consecutive exons in close proximity and helps to ensure the correct order of exons during splicing. **c** | Alternative splicing and alternative polyadenylation (APA): packaging of precursor mRNA (pre-mRNA) could either hide strong splice sites or alternative exons to promote exon skipping, or could expose weak splice sites to facilitate exon inclusion during alternative splicing. Packaging of nascent transcripts could also hide non-canonical weak polyadenylation signals (PASs) or inappropriate PASs that are present in introns. This way, premature cleavage and polyadenylation (PCPA) of transcripts or of 3′ untranslated regions (UTRs) that are too short could be prevented. Packaging could also expose alternative PASs to facilitate APA. **d** | mRNA export: packaging of pre-mRNA can help to estimate the length of transcribed RNAs. Heterogeneous nuclear ribonucleoprotein C (hnRNPC) acts as a 'molecular ruler' and sorts RNA polymerase II (Pol II) transcripts before nuclear export. Sorting is based on a competition between hnRNPC and phosphorylated adaptor for RNA export (PHAX) for binding to the nuclear cap-binding complex (CBC; competition shown by 'T' bar). The hnRNPC tetramer cannot stably bind short transcripts (of <200 nt; right-hand panel), allowing PHAX to be recruited to the CBC and committing the RNA to the small nuclear RNA (snRNA) export pathway by chromosomal region maintenance protein 1 (CRM1). When a nascent transcript bound by the CBC becomes longer than 200 nt, hnRNPC binds and inhibits the recruitment of PHAX to the CBC (left-hand panel), committing the transcript to the nuclear RNA export factor 1 (NXF1) mRNA export pathway[39]. m7G, 7-methylguanosine.

within introns — suggesting the formation of basic units of RNA packaging[20]. Additional crosslinked uridines that were ~300 nt apart also suggested that hnRNPC tetramers can oligomerize to form higher-order structures[20]. On the basis of these findings, it was proposed that hnRNPC could indeed act as an 'RNA nucleosome' that incorporates long regions of nuclear pre-mRNA in several repeating units; this could serve several functions during nuclear steps of RNA processing and export. For example, RNA wrapping could bring distant regions of nascent RNA into close proximity, such as constitutive or alternative exons. Wrapping could also hide non-canonical and cryptic polyadenylation signals (PASs) to avoid premature cleavage and polyadenylation (PCA), or it could hide inappropriate splicing signals[31] (FIG. 2 and see below).

Packaging and co-transcriptional splicing. Nuclear hnRNP particles must accommodate the spliceosome, which is a large RNA–protein machine the size of the ribosome. Because the spliceosome assembles co-transcriptionally on pre-mRNA[17,32], because hnRNPC binds co-transcriptionally[20], and because nascent RNPs contain spliceosomal components as well as hnRNP proteins[20,33], it was proposed that pre-mRNA is co-transcriptionally packaged. Electron microscopy images of nascent transcripts containing spliceosomal short nuclear RNPs (snRNPs) that lie close to the DNA axis support this concept[34]. However, it is currently unclear whether packaging occurs by an orderly mechanism to produce uniform structures of a defined composition.

Co-transcriptional packaging can have several functions: preventing long, linear pre-mRNA molecules from physical entanglement with DNA during transcription[34], which could lead to the formation of R loops[35]; holding splicing intermediates in close proximity to ensure the correct order of events[36]; or hiding strong splice sites or exposing weak splice sites for alternative splicing[20] (FIG. 2). Interestingly, several hnRNP proteins, such as hnRNPA1, hnRNPC and PTB, are regulators of alternative splicing[37]. A recent single-molecule study showed that co-transcriptional splicing can be blocked *in vivo* when the polypyrimidine tract near the 3′ end of introns is sequestered by PTB binding[38]. Thus, co-transcriptional pre-mRNA packaging also regulates alternative splicing.

Packaging and mRNA export. Co-transcriptional packaging is also important for mRNA export (BOX 2). It was recently shown that hnRNPC sorts RNA polymerase II (Pol II) transcripts into two distinct classes of export-competent RNPs: long mRNAs, which are

Premature cleavage and polyadenylation
(PCA). Misprocessing of precursor mRNAs (pre-mRNA) by the cleavage and polyadenylation machinery. Truncated transcripts arise through the use of cryptic or inappropriate polyadenylation signals present at the 5′ end or within introns of the pre-mRNA.

R loops
Hybrid structures consisting of RNA and DNA in which RNA displaces a DNA strand to hybridize to its complementary DNA sequence.

Box 2 | **mRNP export and remodelling**

In eukaryotes, precursor mRNA (pre-mRNA) processing is spatially and temporally separated from mRNA translation by the nuclear membrane. Nuclear pores perforate the nuclear envelope to allow nucleocytoplasmic transport of export-competent ribonucleoprotein particles (RNPs) by specific export receptors. Mature messenger RNPs (mRNPs) are exported to the cytoplasm by a heterodimeric mRNA export receptor consisting of nuclear export factor 1 (NXF1) and NTF2-related protein 1 (NXT1). Through interaction with different adaptor proteins, such as SR proteins, ALYREF, UAP56-interacting factor (UIF), cleavage and polyadenylation specificity factor subunit 6 (CPSF6), UAP56 and germinal centre-associated nuclear protein (GANP)[45] (FIG. 4), NXF1–NXT1 can couple sequential RNA maturation steps and export different mRNA classes[45]. NXF1 is thought to escort mRNA adaptors and their mRNA cargo through the nuclear pore complex (NPC). After they have entered the cytoplasm, adaptor proteins are released from NXF1 and might remain bound to the exported mRNA cargo to fulfil cytoplasmic functions.

What if mRNPs are too large to exit through nuclear pores? It was recently shown in *Drosophila melanogaster* neurons that large RNP granules (of 0.2 μm diameter) containing polyadenylated mRNAs bound to the PABPN1 homologue PABP2 exit the nucleus by budding through the inner and the outer nuclear membranes in a mechanism analogous to budding of herpesviruses[116]. This process, termed nuclear envelope budding (FIG. 4), involves phosphorylation of A-type lamin proteins and is necessary for the proper localization and local translation of several postsynaptic transcripts[116]. These findings uncover an alternative mechanism by which cellular mRNAs can exit the nucleus. Although the molecular mechanisms that drive and regulate nuclear envelope budding are unclear and the protein composition of these large mRNPs is largely unknown, this example also demonstrates the importance of mRNP size.

exported by nuclear RNA export factor 1 (NXF1); and short (<200 nt) snRNAs, which are exported by the chromosomal region maintenance protein 1 (CRM1; also known as XPO1)–RAN·GTP pathway (FIG. 2). hnRNPC carries out this task by measuring the length of newly transcribed RNAs as they protrude from RNA Pol II[39]. This 'molecular ruler' activity of hnRNPC is due to its capacity to wrap long stretches of RNA around tetramers[39]. It was shown that hnRNPC can bind only when a nascent transcript reaches a length of 150 to 250 nt, and this binding subsequently inhibits RNA export by the CRM1–RAN·GTP pathway[39]. Although hnRNPC is mostly bound to introns[20] and is therefore largely absent from spliced mRNAs[40], residual bound hnRNPC must be removed from transcripts before export because it was shown not to shuttle to the cytoplasm[40].

The above model implies that all protein-coding transcripts must be bound by hnRNPC to be sorted. However, hnRNPC binding was mapped to only 55% of all protein-coding genes by iCLIP[20]. This discrepancy might be explained by the fact that not all mRNAs are expressed in HeLa cells and that sequencing in iCLIP studies is often not sufficiently deep to detect all bound RNAs above a stringent cut-off; therefore, many true binding sites might be missed[20]. Whether concepts such as functional packaging generally apply will have to be addressed in future studies.

Nuclear mRNPs

The proteomic repertoire of nuclear mRNPs. Little is known regarding the complete protein content of any individual mRNP in any cellular compartment. So far, comprehensive proteomes have been determined for only two discrete mRNPs, which were affinity-purified from HeLa cell nuclear extracts following completion of *in vitro* splicing[40]. Forty-five proteins were identified, including many hnRNP proteins, indicating that these mRNPs represented bona fide nuclear mRNPs. The assembly of the exon junction complex (EJC) and most nuclear export factors depended on splicing but not on the presence of the CBC. Moreover, several splicing factors, such as SR proteins, remained bound to the spliced RNAs[40], which is consistent with the notion that successive mRNA-processing steps are intimately coupled to one another through shared regulatory RBPs[6] (BOX 3; FIG. 3).

Coupling of splicing to mRNA export. Consistent with the presence of EJC and SR proteins in nuclear post-splicing mRNPs[40,41], it was recently proposed that EJC and SR proteins function together in mRNP packaging for efficient nuclear export, thus linking splicing to mRNA export[42] (FIG. 3). A genome-wide study based on nuclease protection and quantitative mass spectrometry[42], as well as a related CLIP study[43], showed that pre-mRNAs containing several introns accumulate several EJCs after splicing *in vivo*[42,43]. Multimerization among EJCs and numerous SR proteins bound to the same transcript can lead to the formation of higher-order mRNP structures, which may drive overall mRNP compaction[42]. In agreement with this model, it was shown that the Balbiani ring mRNPs of *Chironomus tentans* salivary glands adopt a highly compacted ring-like structure containing several different SR proteins[44].

Some SR proteins also serve as mRNA adaptors that direct bound transcripts to the NXF1 pathway for mRNA export[45] (FIG. 3). Most SR proteins continuously shuttle between the nucleus and the cytoplasm, albeit with different shuttling capacities[46–48]. The shuttling efficiency of individual SR proteins correlates with their RNA-binding capacity, and shuttling of serine/arginine-rich splicing factor 1 (SRSF1) requires both RRM domains to be intact[49,50], arguing for an active role of SR proteins in the export of their mRNA cargo. However, export of mRNAs by SR proteins has so far been shown only for intronless histone H2A reporter transcripts[51,52], which are not spliced.

Another adaptor for NXF1-dependent mRNA export is the transcription export complex 1 (TREX1), which is a multi-protein complex that in yeast links transcription with mRNA export[53]. Although the TREX1 complex is evolutionarily conserved, in humans the efficient recruitment of TREX1 to pre-mRNA is mostly splicing-dependent[54]. A recent study demonstrated that TREX1 releases spliced mRNAs from the site of pre-mRNA splicing for export to the cytoplasm, thereby coupling splicing and mRNA export[55].

Coupling of transcription to mRNA export. Several recent studies have provided surprising examples of RBPs that also connect non-sequential steps in RNA metabolism. For example, although the deposition of TREX1 on mRNA normally occurs during splicing in mammals[54], TREX1 also participates in the nuclear export of several intronless transcripts independently

Exon junction complex
(EJC). A protein complex that is deposited ~24 nucleotides upstream of the exon–exon junctions of newly synthesized, spliced mRNAs. The EJC contains four core proteins — eukaryotic initiation factor 4AIII (EIF4AIII), Y14, mago nashi homologue (MAGOH) and Barentsz (BTZ) — and several loosely associated proteins.

SR proteins
Evolutionarily conserved RNA-binding proteins with essential functions in precursor mRNA (pre-mRNA) splicing in metazoans. Individual SR proteins have distinct RNA-binding capacities and are important regulators of alternative splicing, while some also function in post-splicing steps of gene expression.

Balbiani ring
Chromosome puffs or large diffused uncoiled regions, which are the sites of RNA transcription, in the giant polytene chromosomes of *Chironomus tentans* salivary gland cells.

Box 3 | **Cap-binding complex: connecting nuclear RNA metabolic steps**

The acquisition of the 7-methylguanosine (m^7G) cap at the mRNA 5′ end is the earliest mRNA processing event, starting as soon as nascent transcripts emerge from RNA polymerase II (Pol II) during promoter-proximal pausing induced by DRB-sensitivity-inducing factor (DSIF) and negative elongation factor (NELF)[117]. The m^7G cap structure is immediately bound by the nuclear cap-binding complex (CBC), which is one of the key factors for coupling transcription to subsequent mRNA maturation steps (FIG. 1). The CBC consists of two main subunits: CBP20 and CBP80. Additional ribonucleoprotein particles (RBPs) that interact with the CBC have recently been identified, broadening the functional catalogue of CBC. For example, arsenite-resistance protein 2 (ARS2) is an accessory component of the CBC that co-transcriptionally binds to RNA substrates containing stem–loop structures, such as primary microRNAs and histone transcripts, and controls their transcription, 3′ end processing and stability[118,119]. In yeast, it was shown that CBC interacts with modifier of transcription 1 (Mot1) *in vivo* and stimulates the formation of pre-initiation complexes (PICs) at several promoters through recruitment of Mot1 (REF. 120), showing that an RBP can have a reciprocal synergistic effect on transcription initiation. In human cells, NELF interacts with the CBC and recruits the stem–loop-binding protein (SLBP) to the 3′ end of replication-dependent histone precursor mRNAs. SLBP binds to a highly conserved stem–loop and, assisted by NELF, directs site-specific cleavage to generate the 3′ ends of histone transcripts lacking poly(A) tails[63,121]. As SLBP is also required for the export of histone transcripts to the cytoplasm, CBC links 5′ end capping to mRNA export[62] (FIG. 4). CBC was further shown to interact with several transcription export complex 1 (TREX1) components, this way recruiting export factors to the 5′ end of intronless transcripts and connecting 5′ end capping to mRNA export[60,61] (FIG. 4). Another link between 5′ end capping and mRNA export is the co-transcriptional mRNA sorting by heterogeneous nuclear RNP C (hnRNPC) before mRNP export, in which a competition between hnRNPC and phosphorylated adaptor for RNA export (PHAX) for binding to CBC provides the molecular basis[39] (FIGS 2,4). Finally, CBC is essential for nonsense-mediated mRNA decay (NMD), and it co-purifies with the nuclear exosome targeting (NEXT) complex, thus linking transcription to nuclear surveillance[79,122].

of splicing[56,57]. Efficient export requires the presence of GC-rich export-promoting sequences at the 5′ end of these transcripts[57,58]. It was proposed that such sequences, which are termed signal sequence coding regions (SSCRs[58]) or cytoplasmic accumulation regions (CARs[57]), serve as mRNA identity elements[59] that distinguish bona fide mRNAs from similarly sized transcripts containing random sequences and mediate their export[55,57]. Recruitment of the TREX1 components THOC5, ALYREF (also known as THOC4) and UAP56 (also known as DDX39B) to the 5′ end of intronless mRNAs occurs through interaction with the cap-binding complex (CBC)[60,61] (BOX 3; FIG. 4). It was further shown that ALYREF and THOC5 are required for the export of intronless heat-shock protein 70 (*HSP70*) mRNA *in vivo*[56]. This suggests that, similarly to in yeast, in metazoans TREX1 functions in co-transcriptional loading of mRNA export factors to a subset of nascent transcripts, thereby connecting transcription to mRNA export[56]. This alternative mRNA export pathway, which is termed alternative RNA export (ALREX), requires NXF1 and the CBC[59] and constitutes a novel splicing-independent mechanism to export intronless mRNAs (FIG. 4). By contrast, replication-dependent histone mRNAs, which also lack introns, are exported to the cytoplasm by the stem–loop-binding protein (SLBP)[62], which is recruited to the 3′ end of histone transcripts through interaction with the CBC and negative elongation factor (NELF)[63]. This export pathway does not require NXF1 (BOX 3; FIG. 4).

A newly identified mammalian TREX2 complex, originally characterized in yeast, also couples transcription and mRNA export by facilitating the passage of mature mRNPs from transcription foci to the nuclear membrane[64]. This step is important, as most active transcription sites in metazoan cells are concentrated deep within the nucleus. TREX2 has four subunits: germinal centre-associated nuclear protein (GANP; also known as MCM3AP), PCI domain-containing protein 2 (PCID2), the 26S proteasome complex subunit DSS1, and enhancer of yellow 2 transcription factor (ENY2). All four might participate in the export of bulk polyadenylated RNA[65], but so far only GANP has been shown to interact with NXF1 and to chaperone mature mRNPs to nuclear pores[64] (FIG. 4). It is still unclear whether TREX2 activity is dependent on splicing.

Coupling of alternative polyadenylation to mRNA export. Cleavage and polyadenylation (CPA) factors piggyback on Pol II from the start of transcription and are thereby efficiently delivered to the polyadenylation sites[66] (FIGS 1,3). The coupling of 3′ end processing and mRNA export ensures that Pol II transcription correctly terminates before the mRNP is released from the transcription unit and that mature mRNA is packaged together with poly(A)-binding proteins (PABPs) and mRNA export factors[66]. Alternative polyadenylation (APA), which generates multiple transcripts with different 3′ untranslated regions (3′UTRs), is normally linked to splicing[66,67], but little is known of how APA is linked to mRNA export. Interestingly, a splicing-independent role of the U1 snRNP was recently discovered in the suppression of APA[68]. It was shown that U1 snRNP binding protects pre-mRNAs from premature cleavage and polyadenylation (PCPA) at cryptic PASs scattered throughout introns or at non-canonical PASs in 3′UTRs[68,69] (FIG. 3). Binding to PAS regions probably occurs through the U1 snRNP component U1-70K[69]. Because the abundance of U1 snRNP in the cell seems to determine the length of the cleaved pre-mRNAs, the U1 snRNP could also be regarded as a 'molecular ruler'[68].

The nuclear poly(A)-binding protein PABPN1 acts as another 'molecular ruler' by binding to poly(A) tails of transcripts during polyadenylation and measuring their length[70]. Whereas one PABPN1 protein requires a minimum of ~12 As to bind, electron microscopic studies have shown that many PABPN1 proteins coating poly(A) tails form spherical particles with a defined size, which accommodate ~250 nt[71]. PABPN1 is also required for general mRNA export and shuttles to the cytoplasm[72]. Furthermore, it was recently demonstrated that PABPN1 directly interacts with weak and non-canonical PAS sequences *in vivo* and efficiently competes with the cleavage and polyadenylation machinery. Depletion of PABPN1 resulted in a genome-wide shortening of 3′UTRs[73]. Several important findings came from this study: it revealed an unexpected function for PABPN1 in the regulation of APA, it suggested that PABPN1 is already bound to nascent RNA before cleavage, and it indicated that PABPN1 can bind to hexanucleotide PAS

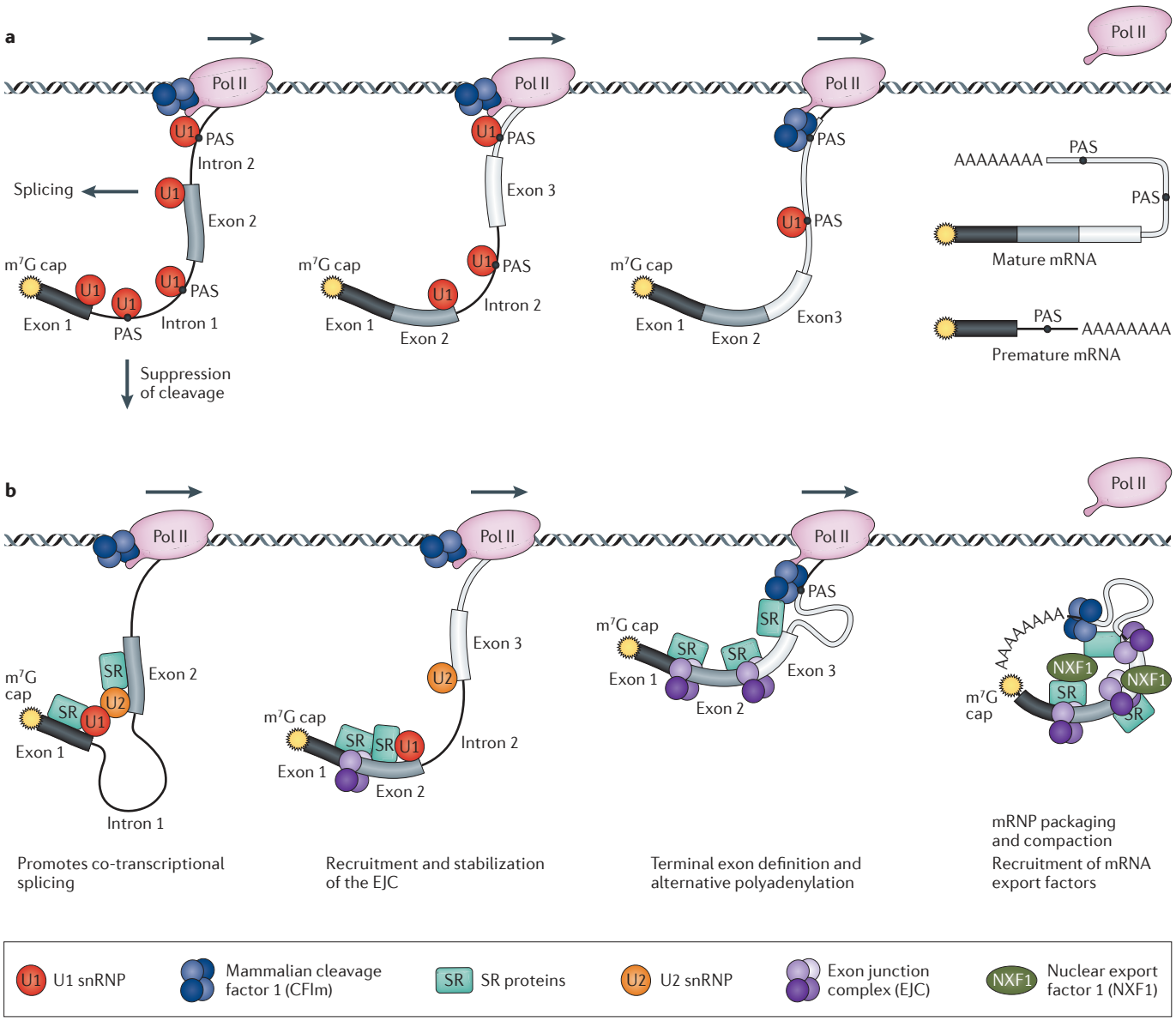

Figure 3 | **Coupling of sequential and non-sequential steps in RNA metabolism by RBPs.** **a** | U1 small nuclear ribonucleoprotein particles (snRNPs) couple splicing and alternative polyadenylation. This is an example in which copies of the same factor (but not the same molecule) carry out two distinct functions in RNA processing. U1 snRNP is essential for splicing and binds to 5′ splice sites (5′ss) of nascent transcripts. Recently, it was also found to bind close to cryptic or inappropriate poly(A) signals (PASs) within introns, thereby protecting long intron-containing transcripts from premature cleavage and polyadenylation (PCPA)[68,69]. Both processes probably occur independently of each other and are not carried out by the same U1 snRNP molecule, as binding to the 5′ss occurs through RNA–RNA pairing between mRNA and U1 snRNA, whereas binding to the PAS region probably involves the U1 snRNP protein U1-70k. The PAS region comprises the highly conserved hexanucleotide AAUAAA found 10–30 nucleotides (nt) upstream of the cleavage site and a less conserved U-rich or GU-rich element located downstream of the cleavage site. **b** | SR proteins couple co-transcriptional splicing to mRNA export. This is an example in which the same bound RNA-binding protein (RBP) molecule participates in several sequential RNA-processing steps. SR proteins bind to specific binding sites within nascent transcripts and aid in the recruitment of the U1 and U2 snRNPs, thereby promoting co-transcriptional splicing. Remaining bound to spliced exons, they subsequently promote the recruitment of the exon junction complex (EJC) and help to stabilize its binding to the transcript[42,43]. Through alternative splicing of terminal exons or interaction with cleavage and polyadenylation factors (for example, mammalian cleavage factor I (CFIm)), SR proteins can also influence the choice of poly(A) sites. As they remain bound to mature mRNAs, SR proteins help in the packaging and compaction of messenger RNPs (mRNPs) through tight interactions with EJC components and other SR proteins bound to the same transcript[42]. Ultimately, they recruit nuclear RNA export factor 1 (NXF1) to the compact, mature mRNP to allow efficient nuclear export. m7G, 7-methylguanosine; Pol II, RNA polymerase II.

sequences within the transcript and not only to growing poly(A) tails at the transcript end, as was previously thought[70]. It needs to be determined whether these internally bound PABPN1 proteins that suppress APA promote subsequent mRNA export.

A recent genome-wide CLIP study that mapped the *in vivo* binding sites of all core CPA factors uncovered that cleavage factor CPSF6, which is a subunit of the CPA factor CFIm, is a positive regulator of APA and is the most important factor for the choice of an alternative poly(A) site in a transcript[74]. Interestingly, CPSF6 was shown to shuttle between the nucleus and the cytoplasm, to interact with the mRNA export receptor NXF1 and to stimulate mRNA export. Therefore, CPSF6 could potentially connect APA to mRNA export[75] (FIG. 4).

Coupling of splicing to nuclear surveillance. Co-transcriptional recruitment of RBPs protects the mRNA from degradation through competition with the nuclear RNA surveillance machineries, such as the nuclear exosome[6]. The nuclear exosome recognizes and degrades processing intermediates and improperly formed mRNAs in the nucleus[76]. To distinguish export-competent mRNPs from those that are destined for destruction, the nuclear exosome requires adaptor RBPs. Recent studies in yeast[77,78] provided a comprehensive list of exosome-associated adaptor RBPs. These included: Rrp47 (also known as Lrp1), which is required for 3′ end processing of precursor ribosomal RNAs (pre-rRNAs), snRNAs and snoRNAs; the Trf4–Air2–Mtr4 polyadenylation (TRAMP) complex, which adds oligo(A) tails to substrate RNAs to assist the exosome in degrading highly structured RNAs; the Nrd1–Nab3–Sen1 complex, which co-transcriptionally recruits TRAMP and the exosome to a subset of bound RNA Pol II and RNA Pol III transcripts; and Pab2, which selectively targets intron-containing pre-mRNAs for degradation[72,77].

Although the TRAMP complex and exosome components are evolutionarily conserved in metazoans, nuclear surveillance is poorly understood in metazoans. So far, only two exosome adaptor complexes have been identified, and their nuclear distribution and activities are different from those in yeast[79]. The human TRAMP complex seems to be restricted to the nucleolus and is involved only in the degradation of aberrantly processed rRNAs[79,80]. The other identified exosome adaptor — the nuclear exosome targeting (NEXT) complex — is located in the nucleoplasm and is required for the degradation of promoter upstream non-coding transcripts (PROMPTs)[79]. The NEXT complex contains MTR4 (also known as SKIVL2) and two previously uncharacterized RBPs: RBM7 and zinc finger CCHC domain-containing protein 8 (ZCCHC8). Interestingly, ZCCHC8 was co-purified with CBC components and several members of the SR protein family[79], and RBM7 was previously shown to interact with splicing factors, including SRSF3 (REF. 81). This finding revealed an unexpected connection between splicing and exosome-mediated degradation and suggests that SR proteins and the CBC might assist the NEXT complex in the recognition of aberrantly spliced transcripts. Another connection between splicing and nuclear surveillance was recently revealed: MTR4 and the nuclear exosome component RRP6 interact with several proteins associated with the spliceosomal U4–U5–U6 tri-snRNP complex[82]. Despite these connections, so far none of the known human exosome targets is spliced.

Apart from exosome-mediated RNA degradation, stability of nuclear transcripts can also be affected by alternative splicing through inclusion of premature translation termination codons (PTCs), which activate the nonsense-mediated mRNA decay (NMD) pathway[7]. Initially, NMD was considered only as an RNA surveillance mechanism that removes aberrant or nonsense mRNAs. It is now clear that NMD also has an important role in post-transcriptional regulation of gene expression through specific alternative splicing events, which are termed regulated unproductive splicing[83,84]. Unproductive splicing generally affects the expression of regulatory splicing factors, such as PTB and SR proteins. SRSF2 and SRSF3 have long been known to regulate their own expression through this mechanism[85,86], and a similar feedback regulation has been suggested for all SR protein family members[83,84,87]. Indeed, a recent iCLIP study showed that SRSF4 also regulates its own expression through unproductive splicing *in vivo*[50] and that SRSF3 can cross-regulate the expression of several other SR proteins in a similar manner[50]. Remarkably, it was recently discovered that NMD targets that result from unproductive splicing have more EJC bound than do normally spliced transcripts[42]. As mammalian NMD is highly dependent on the presence of EJCs on the transcripts[88], and as previous reports have shown that some SR proteins can stimulate NMD[89,90], it has been proposed that the EJC helps SR proteins to remain bound to unproductively spliced transcripts after splicing and also that the EJC and SR proteins act together to direct these transcripts to the NMD machinery[42].

Cytoplasmic fate

Coupling transcription and cytoplasmic mRNA decay. It is still not clear where the decay of NMD targets takes place, but owing to the requirement of translation, it is generally assumed that it takes place in the cytoplasm. However, NMD must occur during or immediately after nuclear export, because many NMD substrates are enriched in nuclear preparations[91], providing one example in which the cytoplasmic fate of mRNAs is determined in the nucleus, where the mRNPs assemble.

A direct connection between transcription and cytoplasmic mRNA stability was recently demonstrated in yeast. Several reports convincingly showed that transcriptional regulation and mRNA decay rates are connected[92,93] and that promoters could contribute to mRNA stability control[94,95], thus coupling nuclear transcription to cytoplasmic decay. Interestingly, the mitotic exit network protein Dbf2, a member of the nuclear Dbf2-related (NDR) protein kinase family, binds co-transcriptionally to cell-cycle-regulated transcripts and regulates their decay independently of its kinase function[94]. It will be interesting to study

Exosome
A protein complex that has 3′ to 5′ exonuclease activity (although an endonuclease activity has also been described). Two forms of the exosome have been characterized that differ in their associated cofactors and cellular localization (one is nuclear and one is cytoplasmic).

Promoter upstream non-coding transcripts
(PROMPTs). A recently discovered class of human RNAs. PROMPTs are produced upstream of promoters of active protein-coding genes. They are mainly nuclear and have poly(A) tails and 5′ cap structures.

Nonsense-mediated mRNA decay
(NMD). The process by which mRNAs containing premature termination codons are destroyed to preclude the production of truncated and potentially deleterious protein products. It is also used in combination with specific alternative splicing events to control the levels of some proteins.

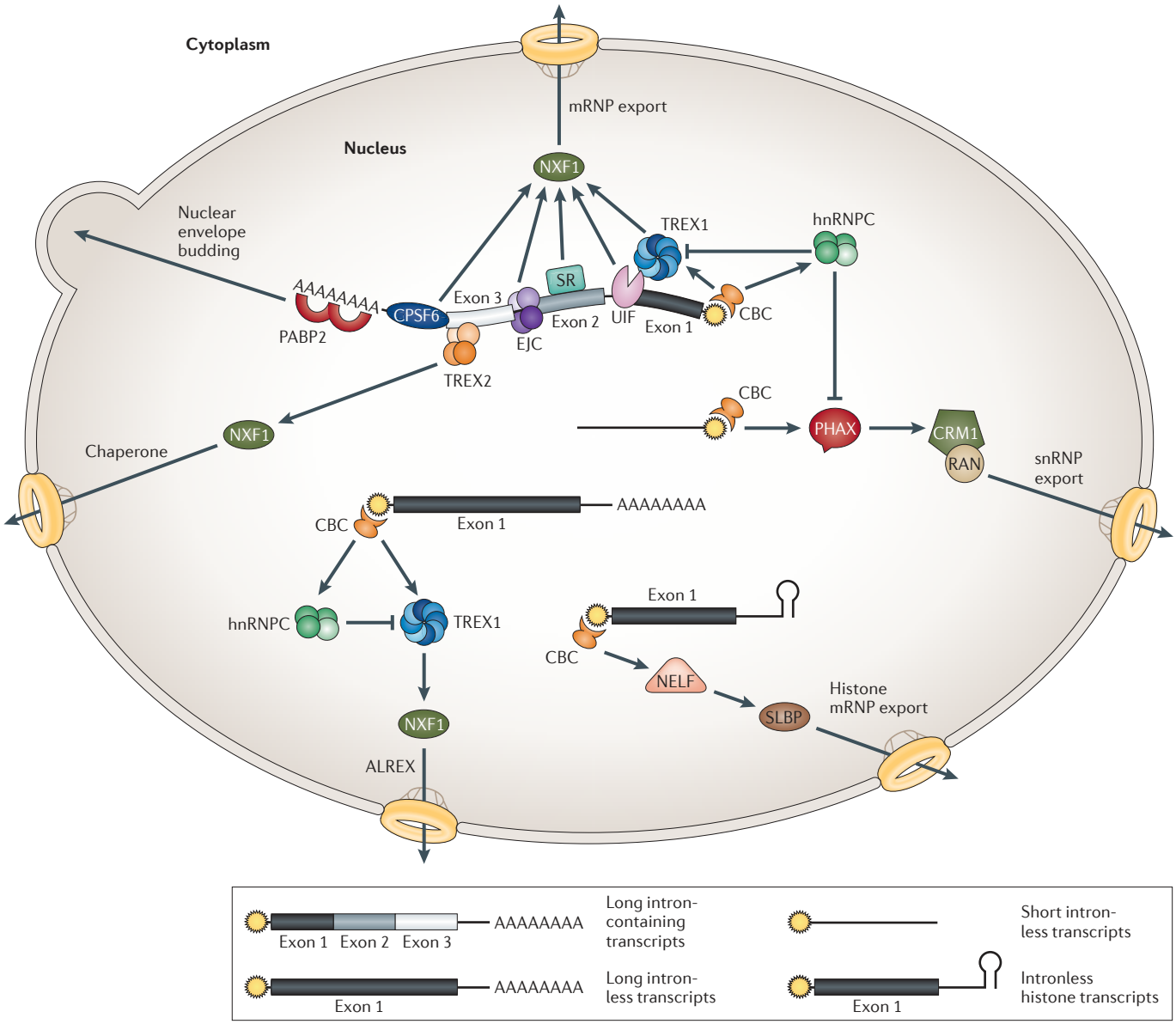

Figure 4 | **Export of mRNPs.** Examples of different export-competent messenger ribonucleoprotein particles (mRNPs) containing four different classes of RNA polymerase II (Pol II) transcripts and their different routes to exit the nucleus. Mature, spliced mRNPs are normally exported by the nuclear RNA export factor 1 (NXF1)-dependent pathway. Through interaction with different mRNA adaptor proteins — including transcription export complex 1 (TREX1)[124], UAP56-interacting factor (UIF)[125], cleavage and polyadenylation specificity factor 6 (CPSF6)[126] or SR proteins[127] — NXF1 is recruited to long intron-containing transcripts in a splicing-dependent manner. When mRNPs are too big to pass through nuclear pores, they may exit through nuclear envelope budding[116]. Replication-dependent intronless histone mRNAs are exported by stem–loop-binding protein (SLBP)[62], which is recruited to these transcripts through interaction with the cap-binding complex (CBC) and negative elongation factor (NELF)[63]. Transcription export complex 2 (TREX2) chaperones mature mRNPs from interior processing foci to the nuclear pores. All four subunits of TREX2 (namely, germinal centre-associated nuclear protein (GANP), PCI domain-containing protein 2 (PCID2), DSS1 and enhancer of yellow 2 transcription factor (ENY2)) participate in mRNA export and probably form a complex, but so far only GANP was shown to interact directly with NXF1 (REFS 64,65). The metazoan TREX1 complex consists of the multisubunit THO complex (THOC1–7), and the proteins UAP56, TEX1 and CIP29 (REF. 124). The export of long intronless transcripts occurs through an alternative RNA export pathway (ALREX) and involves the recruitment of the subunits ALYREF, THOC5 and UAP56 to the 5′ end of the transcript through interaction with the CBC, independently of splicing, which in turn recruit NXF1 to the mRNA[57,124]. Phosphorylated adaptor for RNA export (PHAX), heterogeneous nuclear RNP C (hnRNPC) and ALYREF compete with each other for the interaction with the CBC and binding to the cap-proximal RNA region[39]. Short intronless transcripts, which hnRNPC cannot stably bind, are bound by PHAX instead and are subsequently exported via the chromosomal region maintenance protein 1 (CRM1)–RAN·GTP pathway. EJC, exon junction complex; PABP2, poly(A) binding protein 2 (homologue of nuclear poly(A) binding protein 1).

whether transcription and mRNA decay are also linked in other species and to study how this connection is brought about.

mRNA localization. Localization of mRNAs, which allows the spatial and temporal confinement of protein synthesis to specific sites within the cytoplasm, requires active transport of translationally silenced mRNPs. *Cis*-acting localization sequences within mature mRNAs are recognized by cytoplasmic RBPs, which then recruit translational repressor and motor proteins that translocate the assembled mRNPs to their final destination. For example, in *Saccharomyces cerevisiae*, the RBP She2 binds to specific sequence elements within localized transcripts called 'zipcodes' and directs the localization of bound mRNAs to the bud tip[96]. The fate of localized mRNAs was believed to be determined only after nuclear export and was believed to involve mRNP remodelling. However, recent studies challenge this view. It was discovered that She2 recognizes its target sequences in nascent mRNAs in the nucleus, this way connecting nuclear transcription to mRNA localization in the cytoplasm[97]. She2 first interacts with transcribing Pol II and then 'hops' onto the nascent *ASH1* mRNA as soon as the zipcode is transcribed. Subsequently, the mRNP shuttles to the cytoplasm to recruit the translational repressor proteins Loc1 and pumilio homology domain family member 6 (Puf6)[98] and a type V myosin, Myo4, which are involved in proper mRNA localization[99].

RBPs that bind their mRNA targets in the nucleus and regulate their cytoplasmic localization have also been described in metazoans. For example, the zipcode-binding proteins IGF2BP1 (also known as ZBP1) and ZBP2 regulate the localization of β-actin mRNA to the leading edge of chicken embryo fibroblasts and growth cones in developing neurons[100]. *IGF2BP1*, *ZBP2* and β-actin mRNAs were shown to colocalize in nuclear foci that coincide with active β-actin transcription sites, but it is still unclear whether they directly bind to β-actin mRNAs in the nucleus or whether they are bound to RNA Pol II[101].

Another example is the cytoplasmic localization of *oskar* mRNA to the posterior pole of the *Drosophila melanogaster* oocyte, an event that is dependent on splicing in the nucleus[102]. The basis for this connection was recently revealed[103]. It was shown that splicing of *oskar* serves two important functions: first, it assembles a bipartite *cis*-acting localization element termed SOLE by joining the first two exons together; and, second, splicing loads the EJC onto the transcript, which is required for nuclear export. The EJC binds adjacent to the SOLE and, together with the *oskar* 3′UTR, forms a functional unit that promotes proper kinesin-based translocation of *oskar* mRNPs to the posterior pole[103]. It is currently unclear whether a SOLE-binding factor is recruited in the nucleus or after export in the cytoplasm and whether it acts cooperatively with the EJC for localization.

Conclusions: new concepts of mRNP assembly

A wealth of valuable knowledge about the composition, size and function of mRNPs has been obtained in the past 5 years through a combination of global *in vivo* and *in vitro* studies. The recent introduction of interactome capture has provided a comprehensive list of proteins that bind to RNA in general, revealing that the repertoire of RBPs is far greater than anticipated. Several new RNA-binding domains have been identified, and some of them have been validated *in vivo*[11,12]. RNA-binding profiles of RBPs that contain multiple RBDs or complexes of RBPs have revealed that mRNP packaging is of crucial importance in gene expression and that many RBPs act as molecular rulers that measure or influence the length of mRNAs or that sort RNAs according to their length. Furthermore, recent studies showing that non-sequential steps in RNA metabolism are coupled by several different RBPs have revealed novel and exciting roles for 'old' players, such as the CBC, SR proteins, PABPN1 and hnRNPC.

It has generally been assumed that regulatory RBPs controlling different steps in the mRNA life cycle interact with the mRNA only when that function is needed, predicting that there should be constant changes in the inventory of RBPs associated with an mRNA. The examples discussed here — describing connections between nuclear transcription or splicing and cytoplasmic mRNA localization or decay[93–95,103] — suggest a fundamentally different model in which cellular mRNAs may become 'fully functionally configured' during their synthesis[94,104]. These mRNAs exit the nucleus equipped with all the necessary regulatory RBPs to control their localization, translation and decay in the cytoplasm. These RBPs may be shed from the transcript after completion of each step. However, RBPs generally bind RNA with fairly low affinity, and it is currently unclear how long the same protein remains bound to a given target. Post-translational modifications of RBPs, such as phosphorylation or the binding of other proteins, can increase RBP affinity so that they remain bound to the message[42], or they can decrease affinity to facilitate mRNP remodelling[105]. Reversible modifications of the mRNA itself, such as methylation of adenosine or cytosine, can also affect RBP binding[106,107]. Furthermore, increased competition owing to concentration differences between cellular compartments should also lead to exchange among RBPs. It remains to be seen whether these interesting examples demonstrate novel rules or exceptions to the old rules.

Despite these recent insights, we highlight two outstanding questions: to what extent does the protein composition of an mRNP change during the lifetime of an mRNA; and how many different proteins interact with individual mRNAs? To answer these questions, perhaps the most pressing task for the field is to determine the proteomic content of individual mRNPs. A subsequent challenge will be to understand how the composition of mRNPs is linked to mRNP function and to determine how it changes over time or in different cellular compartments. The availability of techniques that map RNA–protein interaction sites and measure precise quantities of RNA and proteins, including stoichiometric ratios, should ultimately reveal crucial steps in mRNP assembly and remodelling.

1. Fischer, U., Englbrecht, C. & Chari, A. Biogenesis of spliceosomal small nuclear ribonucleoproteins. *Wiley Interdiscip. Rev. RNA* **2**, 718–731 (2011).
2. Watkins, N. J. & Bohnsack, M. T. The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip. Rev. RNA* **3**, 397–414 (2012).
3. Keene, J. D. RNA regulons: coordination of post-transcriptional events. *Nature Rev. Genet.* **8**, 533–543 (2007).
4. Ascano, M., Hafner, M., Cekan, P., Gerstberger, S. & Tuschl, T. Identification of RNA–protein interaction networks using PAR-CLIP. *Wiley Interdiscip. Rev. RNA* **3**, 159–177 (2012).
5. Anko, M. L. & Neugebauer, K. M. RNA–protein interactions *in vivo*: global gets specific. *Trends Biochem. Sci.* **37**, 255–262 (2012).
6. Moore, M. J. & Proudfoot, N. J. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* **136**, 688–700 (2009).
7. Schoenberg, D. R. & Maquat, L. E. Regulation of cytoplasmic mRNA decay. *Nature Rev. Genet.* **13**, 246–259 (2012).
8. König, J., Zarnack, K., Luscombe, N. M. & Ule, J. Protein–RNA interactions: new genomic technologies and perspectives. *Nature Rev. Genet.* **13**, 77–83 (2011).
9. Tutucci, E. & Stutz, F. Keeping mRNPs in check during assembly and nuclear export. *Nature Rev. Mol. Cell Biol.* **12**, 377–384 (2011).
10. Grunwald, D. & Singer, R. H. Multiscale dynamics in nucleocytoplasmic transport. *Curr. Opin. Cell Biol.* **24**, 100–106 (2012).
11. Castello, A. *et al.* Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* **149**, 1393–1406 (2012).
12. Baltz, A. G. *et al.* The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell* **46**, 674–690 (2012).
References 11 and 12 introduce new methods to identify RNA-binding proteins globally in human cell lines, to provide a comprehensive atlas of proteins that can bind to polyadenylated RNAs and to identify novel RNA-binding domains.
13. Mitchell, S. F., Jain, S., She, M. & Parker, R. Global analysis of yeast mRNPs. *Nature Struct. Mol. Biol.* **20**, 127–133 (2013).
14. Kishore, S., Luber, S. & Zavolan, M. Deciphering the role of RNA-binding proteins in the post-transcriptional control of gene expression. *Brief. Funct. Genom.* **9**, 391–404 (2010).
15. Cui, X. A., Zhang, H. & Palazzo, A. F. p180 promotes the ribosome-independent localization of a subset of mRNA to the endoplasmic reticulum. *PLoS Biol.* **10**, e1001336 (2012).
16. Mackereth, C. D. & Sattler, M. Dynamics in multi-domain protein recognition of RNA. *Curr. Opin. Struct. Biol.* **22**, 287–296 (2012).
17. Wahl, M. C., Will, C. L. & Luhrmann, R. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**, 701–718 (2009).
18. MacRae, I. J., Zhou, K. & Doudna, J. A. Structural determinants of RNA recognition and cleavage by Dicer. *Nature Struct. Mol. Biol.* **14**, 934–940 (2007).
19. Lamichhane, R., Solem, A., Black, W. & Rueda, D. Single-molecule FRET of protein–nucleic acid and protein–protein complexes: surface passivation and immobilization. *Methods* **52**, 192–200 (2010).
20. Konig, J. *et al.* iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature Struct. Mol. Biol.* **17**, 909–915 (2010).
Introducing a new method to identify direct binding sites of RNA-binding proteins with a high resolution, termed iCLIP, the authors of this paper indicate that hnRNPC is crucial for the co-transcriptional packaging of most RNAs and provide a functional link to alternative splicing.
21. Kumar, A. & Pederson, T. Comparison of proteins bound to heterogeneous nuclear RNA and messenger RNA in HeLa cells. *J. Mol. Biol.* **96**, 353–365 (1975).
22. Beyer, A. L., Christensen, M. E., Walker, B. W. & LeStourgeon, W. M. Identification and characterization of the packaging proteins of core 40S hnRNP particles. *Cell* **11**, 127–138 (1977).
23. Derman, E., Goldberg, S. & Darnell, J. E. Jr. hnRNA in HeLa cells: distribution of transcript sizes estimated from nascent molecule profile. *Cell* **9**, 465–472 (1976).
24. Mor, A. *et al.* Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells. *Nature Cell Biol.* **12**, 543–552 (2010).
25. Gorlach, M., Burd, C. G., Portman, D. S. & Dreyfuss, G. The hnRNP proteins. *Mol. Biol. Rep.* **18**, 73–78 (1993).
26. Dreyfuss, G., Choi, Y. D. & Adam, S. A. The ribonucleoprotein structures along the pathway of mRNA formation. *Endocr. Res.* **15**, 441–474 (1989).
27. Weighardt, F., Biamonti, G. & Riva, S. The roles of heterogeneous nuclear ribonucleoproteins (hnRNP) in RNA metabolism. *BioEssays* **18**, 747–756 (1996).
28. McAfee, J. G., Soltaninassab, S. R., Lindsay, M. E. & LeStourgeon, W. M. Proteins C1 and C2 of heterogeneous nuclear ribonucleoprotein complexes bind RNA in a highly cooperative fashion: support for their contiguous deposition on pre-mRNA during transcription. *Biochemistry* **35**, 1212–1222 (1996).
29. McAfee, J. G., Shahied-Milam, L., Soltaninassab, S. R. & LeStourgeon, W. M. A major determinant of hnRNP C protein binding to RNA is a novel bZIP-like RNA binding domain. *RNA* **2**, 1139–1152 (1996).
30. Huang, M. *et al.* The C-protein tetramer binds 230 to 240 nucleotides of pre-mRNA and nucleates the assembly of 40S heterogeneous nuclear ribonucleoprotein particles. *Mol. Cell. Biol.* **14**, 518–533 (1994).
31. Neugebauer, K. M. Please hold—the next available exon will be right with you. *Nature Struct. Mol. Biol.* **13**, 385–386 (2006).
32. Bentley, D. L. Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr. Opin. Cell Biol.* **17**, 251–256 (2005).
33. Listerman, I., Sapra, A. K. & Neugebauer, K. M. Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nature Struct. Mol. Biol.* **13**, 815–822 (2006).
34. Wetterberg, I., Zhao, J., Masich, S., Wieslander, L. & Skoglund, U. *In situ* transcription and splicing in the Balbiani ring 3 gene. *EMBO J.* **20**, 2564–2574 (2001).
35. Aguilera, A. & Garcia-Muse, T. R loops: from transcription byproducts to threats to genome stability. *Mol. Cell* **46**, 115–124 (2012).
36. Dye, M. J., Gromak, N. & Proudfoot, N. J. Exon tethering in transcription by RNA polymerase II. *Mol. Cell* **21**, 849–859 (2006).
37. Martinez-Contreras, R. *et al.* hnRNP proteins and splicing control. *Adv. Exp. Med. Biol.* **623**, 123–147 (2007).
38. Vargas, D. Y. *et al.* Single-molecule imaging of transcriptionally coupled and uncoupled splicing. *Cell* **147**, 1054–1065 (2011).
39. McCloskey, A. Taniguchi, I., Shinmyozu, K. & Ohno, M. hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. *Science* **335**, 1643–1646 (2012).
This paper describes a new mechanism in which Pol II transcripts are sorted according to their length prior to nuclear export and identified hnRNPC as the key player.
40. Merz, C., Urlaub, H., Will, C. L. & Luhrmann, R. Protein composition of human mRNPs spliced *in vitro* and differential requirements for mRNP protein recruitment. *RNA* **13**, 116–128 (2007).
41. Anko, M. L., Morales, L., Henry, I., Beyer, A. & Neugebauer, K. M. Global analysis reveals SRp20- and SRp75-specific mRNPs in cycling and neural cells. *Nature Struct. Mol. Biol.* **17**, 962–970 (2010).
42. Singh, G. *et al.* The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* **151**, 750–764 (2012).
This paper presents compelling evidence suggesting that the EJC and SR proteins cooperate in the packaging and compaction of mature mRNPs for efficient nuclear export.
43. Sauliere, J. *et al.* CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. *Nature Struct. Mol. Biol.* **19**, 1124–1131 (2012).
44. Bjork, P. *et al.* Specific combinations of SR proteins associate with single pre-messenger RNAs *in vivo* and contribute different functions. *J. Cell Biol.* **184**, 555–568 (2009).
45. Walsh, M. J., Hautbergue, G. M. & Wilson, S. A. Structure and function of mRNA export adaptors. *Biochem. Soc. Trans.* **38**, 232–236 (2010).
46. Sapra, A. K. *et al.* SR protein family members display diverse activities in the formation of nascent and mature mRNPs *in vivo*. *Mol. Cell* **34**, 179–190 (2009).
47. Lin, S., Xiao, R., Sun, P., Xu, X. & Fu, X. D. Dephosphorylation-dependent sorting of SR splicing factors during mRNP maturation. *Mol. Cell* **20**, 413–425 (2005).
48. Caceres, J. F., Screaton, G. R. & Krainer, A. R. A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. *Genes Dev.* **12**, 55–66 (1998).
49. Delestienne, N. *et al.* The splicing factor ASF/SF2 is associated with TIA-1-related/TIA-1-containing ribonucleoproteic complexes and contributes to post-transcriptional repression of gene expression. *FEBS J.* **277**, 2496–2514 (2010).
50. Anko, M. L. *et al.* The RNA-binding landscapes of two SR proteins reveal unique functions and binding to diverse RNA classes. *Genome Biol.* **13**, R17 (2012).
51. Huang, Y. & Steitz, J. A. SRprises along a messenger's journey. *Mol. Cell* **17**, 613–615 (2005).
52. Erkmann, J. A., Sanchez, R., Treichel, N., Marzluff, W. F. & Kutay, U. Nuclear export of metazoan replication-dependent histone mRNAs is dependent on RNA length and is mediated by TAP. *RNA* **11**, 45–58 (2005).
53. Strasser, K. *et al.* TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* **417**, 304–308 (2002).
54. Reed, R. & Cheng, H. TREX, SR proteins and export of mRNA. *Curr. Opin. Cell Biol.* **17**, 269–273 (2005).
55. Dias, A. P., Dufu, K., Lei, H. & Reed, R. A role for TREX components in the release of spliced mRNA from nuclear speckle domains. *Nature Commun.* **1**, 97 (2010).
56. Katahira, J., Inoue, H., Hurt, E. & Yoneda, Y. Adaptor Aly and co-adaptor Thoc5 function in the Tap-p15-mediated nuclear export of HSP70 mRNA. *EMBO J.* **28**, 556–567 (2009).
57. Lei, H., Dias, A. P. & 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* **108**, 17985–17990 (2011).
58. Palazzo, A. F. *et al.* The signal sequence coding region promotes nuclear export of mRNA. *PLoS Biol.* **5**, e322 (2007).
59. Palazzo, A. F. & Akef, A. Nuclear export as a key arbiter of 'mRNA identity' in eukaryotes. *Biochim. Biophys. Acta* **1819**, 566–577 (2012).
60. Cheng, H. *et al.* Human mRNA export machinery recruited to the 5′ end of mRNA. *Cell* **127**, 1389–1400 (2006).
61. Nojima, T., Hirose, T., Kimura, H. & Hagiwara, M. The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. *J. Biol. Chem.* **282**, 15645–15651 (2007).
62. Sullivan, K. D., Mullen, T. E., Marzluff, W. F. & Wagner, E. J. Knockdown of SLBP results in nuclear retention of histone mRNA. *RNA* **15**, 459–472 (2009).
63. Narita, T. *et al.* NELF interacts with CBC and participates in 3′ end processing of replication-dependent histone mRNAs. *Mol. Cell* **26**, 349–365 (2007).
64. Wickramasinghe, V. O. *et al.* mRNA export from mammalian cell nuclei is dependent on GANP. *Curr. Biol.* **20**, 25–31 (2010).
65. Jani, D. *et al.* Functional and structural characterization of the mammalian TREX-2 complex that links transcription with nuclear messenger RNA export. *Nucleic Acids Res.* **40**, 4562–4573 (2012).
66. Proudfoot, N. J. Ending the message: poly(A) signals then and now. *Genes Dev.* **25**, 1770–1782 (2011).
67. Lou, H., Neugebauer, K. M., Gagel, R. F. & Berget, S. M. Regulation of alternative polyadenylation by U1 snRNPs and SRp20. *Mol. Cell. Biol.* **18**, 4977–4985 (1998).
68. Berg, M. G. *et al.* U1 snRNP determines mRNA length and regulates isoform expression. *Cell* **150**, 53–64 (2012).
This paper establishes U1 snRNP as a molecular ruler and describes how U1 snRNP levels influence the length of transcripts through suppression of PCPA. The authors demonstrate the physiological importance of this role of U1 snRNP in activated neurons.
69. Kaida, D. *et al.* U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468**, 664–668 (2010).
In this study, the authors discovered a new splicing-independent function of U1 snRNP as a suppressor of alternative polyadenylation. The authors show that U1 snRNP protects pre-mRNAs from PCPA by binding to inappropriate poly(A) sites present within introns of pre-mRNAs.
70. Eckmann, C. R., Rammelt, C. & Wahle, E. Control of poly(A) tail length. *Wiley Interdiscip. Rev. RNA* **2**, 348–361 (2011).

71. Keller, R. W. *et al.* The nuclear poly(A) binding protein, PABP2, forms an oligomeric particle covering the length of the poly(A) tail. *J. Mol. Biol.* **297**, 569–583 (2000).
72. Lemay, J.-F., Lemieux, C., St-André, O. & Bachand, F. Crossing the borders: poly(A)-binding proteins working on both sides of the fence. *RNA Biol.* **7**, 291–295 (2010).
73. Jenal, M. *et al.* The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* **149**, 538–553 (2012).
This work identified the nuclear poly(A)-binding protein PABPN1 as a potent suppressor of alternative polyadenylation and revealed the importance of APA in the pathology of a human disease caused by changes in PABPN1 levels.
74. Martin, G., Gruber, A. R., Keller, W. & Zavolan, M. Genome-wide analysis of pre-mRNA 3′ end processing reveals a decisive role of human cleavage factor I in the regulation of 3′ UTR length. *Cell Rep.* **1**, 753–763 (2012).
75. Ruepp, M. D. Schümperli, D. & Barabino, S.M. mRNA 3′ end processing and more—multiple functions of mammalian cleavage factor I-68. *Wiley Interdiscip. Rev. RNA* **2**, 79–91 (2011).
76. Lykke-Andersen, S., Brodersen, D. E. & Jensen, T. H. Origins and activities of the eukaryotic exosome. *J. Cell Sci.* **122**, 1487–1494 (2009).
77. Butler, J. S. & Mitchell, P. Rrp6, Rrp47 and cofactors of the nuclear exosome. *Adv. Exp. Med. Biol.* **702**, 91–104 (2010).
78. Lemay, J. F. *et al.* The nuclear poly(A)-binding protein interacts with the exosome to promote synthesis of noncoding small nucleolar RNAs. *Mol. Cell* **37**, 34–45 (2010).
79. Lubas, M. *et al.* Interaction profiling identifies the human nuclear exosome targeting complex. *Mol. Cell* **43**, 624–637 (2011).
80. Shcherbik, N., Wang, M., Lapik, Y. R. Srivastava, L. & Pestov, D. G. Polyadenylation and degradation of incomplete RNA polymerase I transcripts in mammalian cells. *EMBO Rep.* **11**, 106–111 (2010).
81. Guo, T. B. *et al.* Spermatogenetic expression of RNA-binding motif protein 7, a protein that interacts with splicing factors. *J. Androl.* **24**, 204–214 (2003).
82. Nag, A. & Steitz, J. A. Tri-snRNP-associated proteins interact with subunits of the TRAMP and nuclear exosome complexes, linking RNA decay and pre-mRNA splicing. *RNA Biol.* **9**, 334–342 (2012).
83. Ni, J. Z. *et al.* Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev.* **21**, 708–718 (2007).
84. Lareau, L. F., Inada, M., Green, R. E., Wengrod, J. C. & Brenner, S. E. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature* **446**, 926–929 (2007).
85. Sureau, A., Gattoni, R., Dooghe, Y., Stevenin, J. & Soret, J. SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. *EMBO J.* **20**, 1785–1796 (2001).
86. Jumaa, H. & Nielsen, P. J. The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. *EMBO J.* **16**, 5077–5085 (1997).
87. Sun, S., Zhang, Z., Sinha, R., Karni, R. & Krainer, A. R. SF2/ASF autoregulation involves multiple layers of post-transcriptional and translational control. *Nature Struct. Mol. Biol.* **17**, 306–312 (2010).
88. Conti, E. & Izaurralde, E. Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Curr. Opin. Cell Biol.* **17**, 316–325 (2005).
89. Zhang, Z. & Krainer, A. R. Involvement of SR proteins in mRNA surveillance. *Mol. Cell* **16**, 597–607 (2004).
90. Sato, H., Hosoda, N. & Maquat, L. E. Efficiency of the pioneer round of translation affects the cellular site of nonsense-mediated mRNA decay. *Mol. Cell* **29**, 255–262 (2008).
91. Muhlemann, O. & Lykke-Andersen, J. How and where are nonsense mRNAs degraded in mammalian cells? *RNA Biol.* **7**, 28–32 (2010).
92. Sun, M. *et al.* Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. *Genome Res.* **22**, 1350–1359 (2012).
93. Dori-Bachash, M., Shema, E. & Tirosh, I. Coupled evolution of transcription and mRNA degradation. *PLoS Biol.* **9**, e1001106 (2011).
94. Trcek, T., Larson, D. R., Moldon, A., Query, C. C. & Singer, R. H. Single-molecule mRNA decay measurements reveal promoter- regulated mRNA stability in yeast. *Cell* **147**, 1484–1497 (2011).
Using single-cell single molecule techniques, this study demonstrates for the first time a direct link between nuclear transcription and cytoplasmic mRNA stability.
95. Bregman, A. *et al.* Promoter elements regulate cytoplasmic mRNA decay. *Cell* **147**, 1473–1483 (2011).
96. Heym, R. G. & Niessing, D. Principles of mRNA transport in yeast. *Cell. Mol. Life Sci.* **69**, 1843–1853 (2012).
97. Shen, Z., St-Denis, A. & Chartrand, P. Cotranscriptional recruitment of She2p by RNA pol II elongation factor Spt4-Spt5/DSIF promotes mRNA localization to the yeast bud. *Genes Dev.* **24**, 1914–1926 (2010).
98. Shen, Z., Paquin, N., Forget, A. & Chartrand, P. Nuclear shuttling of She2p couples *ASH1* mRNA localization to its translational repression by recruiting Loc1p and Puf6p. *Mol. Biol. Cell* **20**, 2265–2275 (2009).
99. Long, R. M., Gu, W., Lorimer, E., Singer, R. H. & Chartrand, P. She2p is a novel RNA-binding protein that recruits the Myo4p-She3p complex to ASH1 mRNA. *EMBO J.* **19**, 6592–6601 (2000).
100. Gu, W., Pan, F., Zhang, H., Bassell, G. J. & Singer, R. H. A predominantly nuclear protein affecting cytoplasmic localization of β-actin mRNA in fibroblasts and neurons. *J. Cell Biol.* **156**, 41–51 (2002).
101. Pan, F., Huttelmaier, S., Singer, R. H. & Gu, W. ZBP2 facilitates binding of ZBP1 to β-actin mRNA during transcription. *Mol. Cell. Biol.* **27**, 8340–8351 (2007).
102. Hachet, O. & Ephrussi, A. Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization. *Nature* **428**, 959–963 (2004).
103. Ghosh, S., Marchand, V., Gaspar, I. & Ephrussi, A. Control of RNP motility and localization by a splicing-dependent structure in *oskar* mRNA. *Nature Struct. Mol. Biol.* **19**, 441–449 (2012).
This paper provides a mechanistic link between nuclear splicing and localized translation of *oskar* mRNA in the cytoplasm.
104. Trcek, T. & Singer, R. H. The cytoplasmic fate of an mRNP is determined cotranscriptionally: exception or rule? *Genes Dev.* **24**, 1827–1831 (2010).
105. Viphakone, N. *et al.* TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nature Commun.* **3**, 1006 (2012).
106. Squires, J. E. *et al.* Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res.* **40**, 5023–5033 (2012).
107. Dominissini, D. *et al.* Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201–206 (2012).
108. Klug, S. J. & Famulok, M. All you wanted to know about SELEX. *Mol. Biol. Rep.* **20**, 97–107 (1994).
109. Martin, F. Fifteen years of the yeast three-hybrid system: RNA–protein interactions under investigation. *Methods* **58**, 367–375 (2012).
110. Niranjanakumari, S., Lasda, E., Brazas, R. & Garcia-Blanco, M. A. Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions *in vivo*. *Methods* **26**, 182–190 (2002).
111. Keene, J. D., Komisarow, J. M. & Friedersdorf, M. B. RIP-chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nature Protoc.* **1**, 302–307 (2006).
112. Licatalosi, D. D. *et al.* HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464–469 (2008).
113. Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141 (2010).
114. Kishore, S. *et al.* A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. *Nature Methods* **8**, 559–564 (2011).
115. Kudla, G., Granneman, S., Hahn, D., Beggs, J. D. & Tollervey, D. Cross-linking, ligation, and sequencing of hybrids reveals RNA–RNA interactions in yeast. *Proc. Natl Acad. Sci. USA* **108**, 10010–10015 (2011).
116. Speese, S. D. *et al.* Nuclear envelope budding enables large ribonucleoprotein particle export during synaptic Wnt signaling. *Cell* **149**, 832–846 (2012).
117. Wu, C. H. *et al.* NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in *Drosophila*. *Genes Dev.* **17**, 1402–1414 (2003).
118. Gruber, J. J. *et al.* Ars2 links the nuclear cap-binding complex to RNA interference and cell proliferation. *Cell* **138**, 328–339 (2009).
119. Gruber, J. J. *et al.* Ars2 promotes proper replication-dependent histone mRNA 3′ end formation. *Mol. Cell* **45**, 87–98 (2012).
120. Lahudkar, S. *et al.* The mRNA cap-binding complex stimulates the formation of pre-initiation complex at the promoter via its interaction with Mot1p *in vivo*. *Nucleic Acids Res.* **39**, 2188–2209 (2011).
121. Marzluff, W. F., Wagner, E. J. & Duronio, R. J. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nature Rev. Genet.* **9**, 843–854 (2008).
122. Maquat, L. E., Hwang, J., Sato, H. & Tang, Y. CBP80-promoted mRNP rearrangements during the pioneer round of translation, nonsense-mediated mRNA decay, and thereafter. *Cold Spring Harb. Symp. Quant. Biol.* **75**, 127–134 (2010).
123. Visa, N., Izaurralde, E., Ferreira, J., Daneholt, B. & Mattaj, I. W. A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during nuclear export. *J. Cell Biol.* **133**, 5–14 (1996).
124. Katahira, J. mRNA export and the TREX complex. *Biochim. Biophys. Acta* **1819**, 507–513 (2012).
125. Hautbergue, G. M. *et al.* UIF, a new mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. *Curr. Biol.* **19**, 1918–1924 (2009).
126. Ruepp, M. D. *et al.* Mammalian pre-mRNA 3′ end processing factor CF I_m68 functions in mRNA export. *Mol. Biol. Cell* **20**, 5211–5223 (2009).
127. Huang, Y., Gattoni, R., Stévenin, J. & Steitz, J. A. S. R. Splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Mol. Cell* **11**, 837–843 (2003).

Acknowledgments
We thank members of our laboratory, T. Pederson and F. McNicoll for helpful discussions and comments on the manuscript. Work in our laboratory on mRNPs is supported by funding from the Max Planck Society and the German Research Foundation (NE-909/3-1 to K.M.N.), and long-term postdoctoral fellowships from the European Molecular Biology Organization (EMBO) and Fonds de recherche en santé du Québec (FRSQ; to M.M.-M.).

Competing interests statement
The authors declare no competing financial interests.

FURTHER INFORMATION
Karla M. Neugebauer's homepage: http://www.mpi-cbg.de/research/research-groups/karla-neugebauer.html

ALL LINKS ARE ACTIVE IN THE ONLINE PDF