

Review

Evolution and Biological Roles of Alternative 3'UTRs

Christine Mayr^{1,*}

More than half of human genes use alternative cleavage and polyadenylation to generate alternative 3' untranslated region (3'UTR) isoforms. Most efforts have focused on transcriptome-wide mapping of alternative 3'UTRs and on the question of how 3'UTR isoform ratios may be regulated. However, it remains less clear why alternative 3'UTRs have evolved and what biological roles they play. This review summarizes our current knowledge of the functional roles of alternative 3'UTRs, including mRNA localization, mRNA stability, and translational efficiency. Recent work suggests that alternative 3'UTRs may also enable the formation of protein–protein interactions to regulate protein localization or to diversify protein functions. These recent findings open an exciting research direction for the investigation of new biological roles of alternative 3'UTRs.

Evolution of 3'UTR Length and Function

A few years ago, it was found that a large fraction of genes use alternative cleavage and polyadenylation to generate alternative 3'UTRs [1–5]. Initially, most efforts concentrated on establishing sequencing protocols to map alternative 3'UTRs transcriptome wide [6–13] and elucidating how alternative 3'UTR ratios are regulated [14–24], two topics that have been summarized in several recent reviews [25–28]. Here, the functional roles of alternative 3'UTRs are discussed and reasons for why they have evolved are suggested.

It is largely unknown how the biological complexity of organisms is achieved. Most cellular processes are performed by proteins through interactions with other proteins [29,30]. Therefore, when asked what makes humans different from worms, it was initially surprising that the number of protein-encoding genes and the coding region length have remained fairly constant during evolution from worms to humans [31–34]. Intuitively, higher protein diversity would enable more complex biological functions. However, higher protein diversity within a constrained space also increases the number of nonfunctional interactions [35,36]. Therefore, protein concentration and diversity have to be limited to avoid overcrowding. When only a fixed number of proteins are available, biological complexity can still be accomplished through compartmentalization to avoid the coexistence of too many protein types, by the strengthening of specific interactions through cooperativity [35], and by enabling multifunctionality of existing proteins.

Genome size, and, thus, the noncoding part of the genome, has dramatically increased during evolution from worms to humans. The expansion in noncoding sequences includes the 3'UTRs of mRNAs. The number of genes that produce alternative 3'UTRs has doubled and 3'UTR length has increased from a median of 140 nucleotides (nt) in worms to 1200 nt in humans and even to 2300 nt when examining genes that generate alternative 3'UTRs [6,13]. This suggests that more complex organisms have increased post-transcriptional gene regulation mediated by 3'UTR elements. 3'UTRs are well known to control mRNA stability, translational efficiency, and mRNA localization [2,5,37–42]. In addition, 3'UTRs were recently shown to mediate protein–protein interactions [43]. By facilitating alternative protein complex formation, alternative 3'UTRs can

Trends

During animal evolution, the number of protein-encoding genes has remained fairly constant but 3' untranslated region (3'UTR) length and the fraction of genes expressing alternative 3'UTRs have increased substantially.

3'UTRs mediate protein–protein interactions. Thus, alternative 3'UTRs facilitate the formation of alternative protein complexes, which can perform alternative protein functions. This diversifies proteome function without a change in amino acid sequence.

About 15–35% of alternative 3'UTRs have significantly different half-lives, which may contribute to the transcriptome diversity of single cells.

Translation rates of mRNAs with alternative 3'UTRs can be differentially affected by signaling. Whereas one isoform generates basal protein levels, translation of the other is induced by signaling.

Long 3'UTRs seem to be bound by many RNA-binding proteins (RBPs) and may exert their functions within RNA granules.

¹Department of Cancer Biology and Genetics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

*Correspondence: mayrc@mskcc.org (C. Mayr).

Table 1. Differences between Single and Multi-UTR Genes

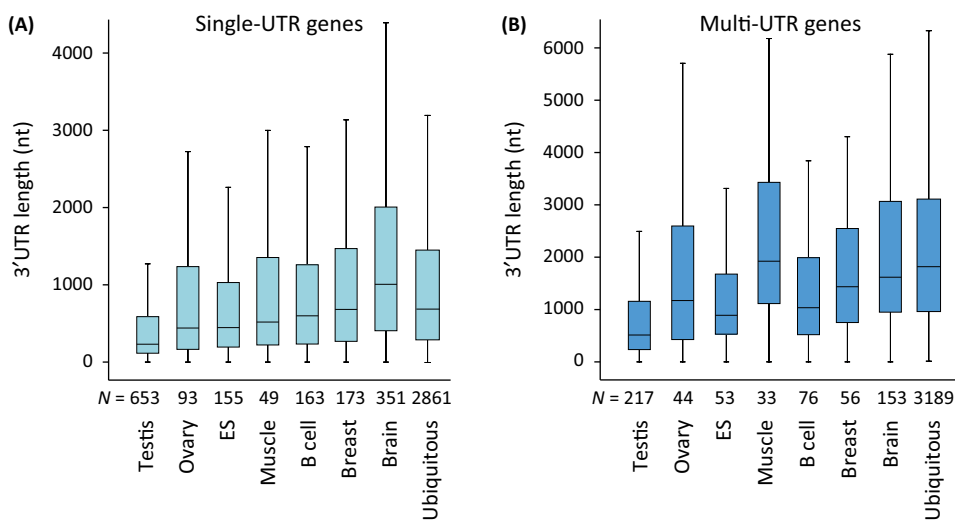
	Single-UTR	Multi-UTR
3'UTR length (nt)	625	2323
Transcription unit length (bp)	20 629	40 519
<i>N</i> expressed in one tissue	1637 (23.7%)	630 (11.1%)
<i>N</i> expressed in two to five tissues	2414 (34.9%)	1854 (32.6%)
<i>N</i> expressed in six or seven tissues	2861 (41.4%)	3189 (56.2%)

diversify protein functions. This may increase biological complexity by implementing multi-functionality of existing proteins.

Single and Multi-UTR Genes Represent Two Distinct Classes of Genes

3'-end sequencing methods have revealed that at least half of human genes generate alternative 3'UTR isoforms [9,12,13]. Transcriptome-wide analyses of alternative 3'UTRs across several normal human tissues and cell lines showed the presence of two classes of genes. One class produces mRNAs with only one 3'UTR, whereas the other class generates alternative 3'UTR isoforms [13]. Single- and multi-UTR genes differ in their genomic architecture, their functions, and their mode of regulation of tissue-specific expression (Table 1).

The median 3'UTR length of single-UTR genes is ~600 nt, but it is ~2300 nt for multi-UTR genes. The increase in 3'UTR length is associated with longer transcription units, which are about twice as long in multi-UTR genes (Table 1). The tissue-specific expression patterns of the two classes of genes also differ significantly. Whereas most genes expressed in only one tissue have single 3'UTRs, more than half of ubiquitously transcribed genes are multi-UTR genes (Table 1) [13]. Furthermore, 3'UTR length is tissue dependent: single- or multi-UTR genes specifically expressed in testis or liver have much shorter 3'UTRs than genes expressed only in the brain (Figure 1) [44–46]. However, ubiquitously transcribed multi-UTR genes have the longest 3'UTRs. They are even longer than the 3'UTRs of brain-specific genes and are about



Trends in Cell Biology

Figure 1. 3' Untranslated Region (3'UTR) Length of Ubiquitously Transcribed and Tissue-Restricted Genes. (A) Single-UTR genes. Here, ubiquitously expressed genes are defined to be expressed in at least six of seven human tissues comprising testis, ovary, embryonic stem (ES) cells, B cells, muscle, breast, and brain. *N*, number of genes in each category. (B) Multi-UTR genes. As in (A).

three times longer than the 3'UTRs of ubiquitously expressed single-UTR genes (Figure 1). This suggests that ubiquitously expressed genes that generate alternative 3'UTRs are especially prone to regulation by elements in their 3'UTRs. The importance of their 3'UTR-based regulation is further supported by higher sequence conservation of multi-UTR genes with longer 3'UTRs [13].

What is the function of widely transcribed genes that generate alternative 3'UTRs? Whereas ubiquitously transcribed single-UTR genes are enriched in genes associated with classical housekeeping functions such as ribosome biogenesis, translation, and energy metabolism, ubiquitous multi-UTR genes are abundant in genes with regulatory functions. They include transcription factors, RNA-binding proteins (RBPs), kinases, phosphatases, and proteins involved in RNA or protein transport [13]. Thus, widely expressed single- or multi-UTR genes encode vastly different classes of proteins. Together with the finding that alternative 3'UTR isoform abundance is highly tissue and cell type specific [9,13], widely expressed regulatory factors may use elements located in their alternative 3'UTRs to achieve tissue-specific expression or function [13,43]. Thus, complex multicellular organisms may use elaborate post-transcriptional regulation to diversify protein functions.

3'UTRs Regulate the Localization, Stability, and Translation of their Cognate mRNAs

mRNAs contain *cis* elements in their 3'UTRs that affect the biology of the mRNA that contains the regulatory elements. The 3'UTR elements, which are bound by RBPs, can control the subcellular localization, the half-life, or the rate of translation of an mRNA. RBPs mediate the various 3'UTR functions through interaction with diverse effector proteins.

Regulation of mRNA Localization

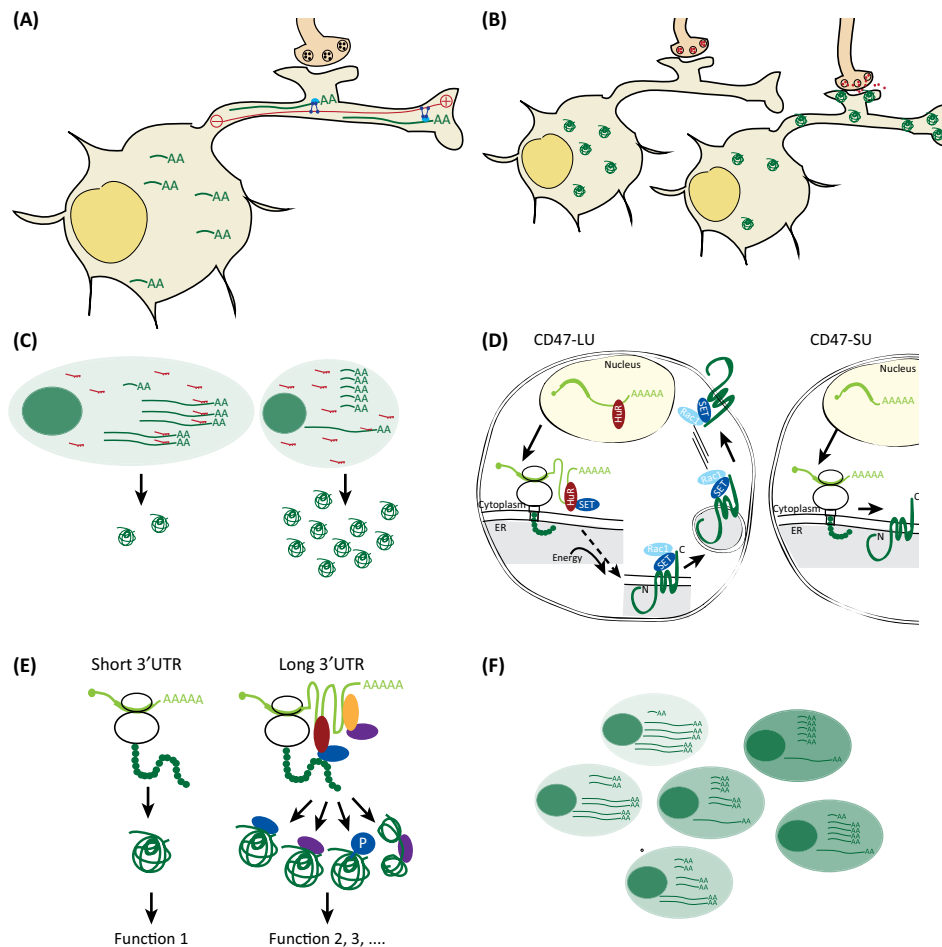
If an RBP that is bound to a 3'UTR element also interacts with a motor protein, it will localize the mRNA to different subcellular compartments [47–49]. Asymmetrically localized mRNAs are used to establish cell polarity, direct asymmetric cell division, or sequester protein activity [37]. In the case of the alternative 3'UTRs of *BDNF*, it was shown that only the long 3'UTR isoform – which contains the localization elements – is localized to dendrites, where it regulates memory formation [38]. *BDNF* is a neurotrophin that plays roles in synaptogenesis and activity-dependent forms of synaptic plasticity [50]. Even modest alterations in *BDNF* levels are associated with behavioral changes in humans and mice, such as abnormal feeding behavior, alterations in episodic memory, and susceptibility to anxiety and depression. Transcription of *BDNF* is regulated by alternative promoters, but, regardless of the promoter, the *BDNF* gene generates alternative 3'UTRs. The short 3'UTR isoform localizes *BDNF* mRNA and protein to the soma of neurons but the long 3'UTR is necessary for localization of *BDNF* to the dendrites (Figure 2A, Key Figure) [38]. Mice that lack the long 3'UTR of *BDNF* have altered dendritic spine morphology and decreased plasticity of dendritic synapses. In summary, mice lacking the long 3'UTR isoform of *BDNF* had impaired long-term potentiation of hippocampal neurons [38]. This study demonstrates that differential localization of *BDNF* 3'UTR isoforms is required for differential functions of *BDNF* in the soma and dendrites of neurons.

Regulation of Translation

If an RBP that is bound to a 3'UTR element interacts with translation initiation factors, it can regulate translation of the mRNA. RBPs that mediate translational repression include Pumilio, Maskin, TIAR, and Musashi-1 [51–53]. As a consequence, alternative 3'UTR isoforms that either include or exclude the *cis* elements can be translated with different efficiencies. This was shown to be the case for the alternative 3'UTR isoforms of *BDNF*, whose translation rates are regulated in a temporal manner. It was shown that the short 3'UTR isoform of *BDNF* is translated in unstimulated hippocampal neurons. Neuronal activation of these neurons led to a switch in

Key Figure

Diverse Biological Roles of Alternative 3' Untranslated Regions (3'UTRs)



Trends in Cell Biology

Figure 2. (A) Regulation of mRNA localization. The long 3'UTR contains a *cis* element that is recognized by an RNA-binding protein (RBP) (light blue). Interaction with a molecular motor (dark blue) enables directional movement along actin fibers (red) and results in localization of the long 3'UTR isoform to the dendrites of neurons. The short 3'UTR isoforms lack the *cis*-acting element and remain localized to the soma. AA indicates a poly(A) tail. (B) Activity-dependent regulation of the translation of alternative 3'UTR isoforms. In the resting state (left) the short 3'UTR isoform of *BDNF* is translated, which results in BDNF protein expression in the soma. After membrane depolarization by synaptic activation (right) the long 3'UTR isoform is translated, which results in BDNF protein expression in dendrites and synapses. (C) Cell type-specific regulation of protein abundance through the interplay of alternative 3'UTRs with miRNAs. Presence of the miRNA and predominant expression of long 3'UTR isoforms in one cell type results in a low protein output whereas predominant expression of short 3'UTR isoforms results in escape from miRNA regulation and leads to a high protein output. AA indicates a poly(A) tail. (D) Alternative 3'UTRs use the scaffold function to regulate 3'UTR-dependent membrane protein localization. HuR-mediated recruitment of SET by the long 3'UTR facilitates the formation of a protein complex containing SET-RAC1-CD47 that enables plasma membrane localization. (E) Alternative 3'UTRs mediate protein-protein interactions to regulate protein function. Long 3'UTRs bind RBPs, which may be able to recruit effector proteins of diverse functions. Effector proteins may act as chaperones to achieve alternative protein folds or as enzymes that add alternative post-translational modifications or, by mediating protein-protein interactions, may enable the formation of alternative protein complexes. (F) Phenotypic diversity of single cells through variability in alternative 3'UTR isoform expression. Each cell of a clonal population expresses varying amounts of alternative 3'UTR isoforms, which may contribute to differences in protein expression and phenotypic diversity among the population. AA indicates a poly(A) tail.

polysome association of the alternative 3'UTR isoforms and resulted in exclusive translation of the long 3'UTR isoform (Figure 2B) [39]. This suggests that the short 3'UTR isoform is responsible for the generation of basal BDNF levels, whereas the long 3'UTR regulates activity-dependent production of BDNF. Neuronal activation also leads to transcriptional upregulation of *BDNF*, but the switch in isoform translation occurs first and enables an immediate increase in BDNF protein levels [39].

Another example of how alternative 3'UTRs regulate translation was shown for *polo*, which encodes Polo-like kinase in *Drosophila* [54]. High levels of Polo are required for the proliferation of abdominal epidermis precursor cells during development. It was demonstrated that the long 3'UTR isoform of *polo* is translated with much higher efficiency than the short 3'UTR isoform. Furthermore, deletion of the proximal polyadenylation signal of *polo* had no phenotypic effect whereas deletion of the distal polyadenylation signal, which abrogated the function of the long 3'UTR isoform, led to death of the flies during development. This result indicated that the higher translation rates of the long 3'UTR isoform are required for the massive proliferation of epidermis precursor cells [54].

Regulation of mRNA Stability

If an RBP that is bound to a 3'UTR element recruits deadenylation or decapping factors, it will destabilize the mRNA and thus affect its half-life. The best-studied *cis* elements that control mRNA stability are miRNA-binding sites and AU-rich elements [40,55–58]. Thus, the binding of RBPs such as AUF1/hnRNP, TTP, or KSRP that bind to AU-rich elements or the miRNA-mediated recruitment of Ago results in mRNA destabilization and short mRNA half-lives. Many proteins whose levels need to be tightly controlled are regulated at the level of mRNA stability, including oncogenes, cytokines, cell cycle regulators, and signaling proteins [57,59–62].

If these genes generate alternative 3'UTRs, the alternative 3'UTR isoforms can control protein abundance levels, which was shown for oncogenes [5]. The shorter 3'UTRs of several oncogenes were more stable than their corresponding long 3'UTR isoforms and produced up to 40-fold more protein [5]. The genes that were chosen for investigation in this study had an enrichment of destabilizing elements in their long 3'UTRs and thus had a high potential for differential regulation of mRNA stability by alternative 3'UTRs [5]. The difference in protein output generated by either the long or short 3'UTR isoform was sufficient for a phenotypic difference, as was shown for *CCND2* and *IGF2BP1* (*IMP-1*) [5]. *IGF2BP1* is an oncogene that when expressed with a short 3'UTR generated IMP-1 protein amounts that were sufficient for oncogenic transformation of fibroblasts in soft agar or in mice [5,63]. By contrast, expression of *IGF2BP1* under the same promoter but with its long 3'UTR had no transforming abilities. Thus, the expression of shorter 3'UTR isoforms of genes whose long 3'UTR isoforms are enriched in negative regulatory elements resulted in increased protein expression. This study also showed that oncogenes can be activated through a change in 3'UTR isoform ratios [5].

This finding led to the prediction that shorter 3'UTRs that escape regulation by 3'UTR elements would be more stable and produce more protein, since most of the 3'UTR elements known at the time, such as miRNA-binding sites, AU-rich, and GU-rich elements, had negative effects on gene expression [40,55,56,64]. When this prediction was tested in studies conducted in yeast, mammalian cell lines, or primary T cells [22,65–67], either no correlation [65,67] or only a weak correlation between shorter 3'UTRs and increased mRNA stability was found [22,66]. Meanwhile, more extensive studies on the effects of regulatory sequences had shown that many 3'UTR elements have positive effects on gene expression [68–70]. Recent data even show that similar numbers of activating and repressive elements are present across all 3'UTRs [69]. These findings confirm the results of transcriptome-wide studies on mRNA stability rates and emphasize that, overall, shorter 3'UTRs are not more stable [22,65–67].

By contrast, the earlier observations that shorter 3'UTRs are more stable were also confirmed in transcriptome-wide studies in which the analysis focused on long 3'UTRs enriched in negative regulatory elements [5,65,66]. Long 3'UTRs that contained predominantly destabilizing elements such as the Puf-binding motif, AU-rich elements, or C-rich elements showed increased decay rates in NIH3T3 cells or yeast [65,66]. In summary, transcriptome-wide studies performed under non-stressed, steady-state conditions found that up to 35% of alternative 3'UTR isoforms differed in their mRNA stability rates, with the shorter 3'UTRs being stable slightly more often [66,68].

Interplay of Alternative 3'UTRs with miRNA-Mediated Regulation Contributes to Cell Type-Specific Gene Expression

Transcriptome-wide analyses of alternative 3'UTR isoforms revealed that the regions that vary between alternative isoforms are enriched in conserved binding sites for RBPs or miRNAs [13,71]. Together with the finding that 3'UTR isoform ratios are highly cell-type specific [9,13], this suggests that cell type-specific gene regulation can be partially accomplished by the interplay of alternative 3'UTR isoform expression and the presence of miRNA-binding sites in distal 3'UTRs. A cell type-specific increase in usage of proximal polyadenylation sites can thus be used to avoid regulation by elements located in distal 3'UTRs. The mRNA of Pax-3, which is an important regulator of myogenesis, is targeted for degradation by miR-206 [72]. However, in a subset of muscle stem cells, high levels of Pax-3 as well as miR-206 are coexpressed. In these cells, high Pax-3 expression is accomplished by the expression of Pax-3 transcripts with shorter 3'UTRs that do not contain the binding site for miR-206. This shows that, under most conditions, Pax-3 is regulated by miR-206, but, in some cell types, expression of the shorter 3'UTR isoform leads to escape from miRNA regulation to enable Pax-3 expression despite expression of the miRNA (Figure 2C) [72]. Avoidance of miRNA-mediated regulation as a means for cell type-specific regulation was confirmed by a transcriptome-wide study on cell lines that showed that approximately 10% of predicted miRNA target sites were affected by the expression of alternative 3'UTRs [73].

The interplay of miRNA expression and alternative 3'UTRs in the regulation of tissue-specific gene regulation is further supported by the finding that ubiquitously transcribed genes that produce only one 3'UTR isoform are enriched in miRNA-binding sites for tissue-specific miRNAs [13]. For these genes, the tissue specificity of the miRNA–target interaction is accomplished by tissue-specific transcription of the miRNAs. By contrast, ubiquitously transcribed genes that generate alternative 3'UTRs are enriched in miRNA-binding sites for ubiquitously expressed miRNAs. Thus, for this class of genes the tissue specificity of the miRNA–target interaction is accomplished by the tissue-specific expression of alternative 3'UTRs [13].

3'UTRs Act as Scaffolds and Mediate Protein–Protein Interactions

The biological functions of 3'UTRs were thought to affect the properties only of the mRNAs that carried the 3'UTR elements. However, a new function of 3'UTRs was discovered that does not alter the fate of the mRNA but instead affects the newly made protein. It was shown that 3'UTRs facilitate the formation of protein–protein interactions. They do so by acting as scaffolds to recruit proteins to the site of translation, which enables the formation of protein complexes with the nascent peptide chain [43]. Protein complex formation can then determine membrane protein localization or protein function.

The scaffold function of 3'UTRs was predominantly studied for the *CD47* gene, which generates alternative 3'UTR isoforms and encodes a membrane protein. It was found that CD47 protein generated by the short *CD47* 3'UTR isoform (CD47-SU) was predominantly retained in the endoplasmic reticulum whereas CD47 protein generated by the long 3'UTR isoform (CD47-LU) efficiently localized to the plasma membrane. CD47-SU and CD47-LU have an identical amino

acid sequence and the difference in protein localization was shown to be independent of mRNA localization, suggesting that the 3'UTRs regulate protein localization [43].

With respect to the mechanism (Figure 2D), it was demonstrated that the long *CD47* 3'UTR binds the RBP HuR, which recruits the effector protein SET [74] to the site of translation. During translation of *CD47*, SET is transferred from the mRNA to the nascent protein, resulting in a protein complex between SET and CD47-LU [43]. The formation of the protein complex depended on the presence of the long 3'UTR of *CD47* because CD47-SU did not interact with SET. SET also binds to RAC1 and active RAC1 translocated SET/CD47-LU to the plasma membrane [43,75]. As a result, CD47-LU efficiently localized to the plasma membrane whereas CD47-SU, which does not have binding sites for HuR and thus does not recruit SET, was retained in the endoplasmic reticulum [43]. Intriguingly, this mechanism of 3'UTR-dependent membrane protein localization was not restricted to the regulation of CD47 but resulted in efficient plasma membrane localization of additional candidates. It was speculated that this localization mechanism may be especially important for membrane proteins with several transmembrane domains, such as G protein-coupled receptors, which are known to be difficult to express in heterologous systems [76]. These results suggest that inclusion of the endogenous 3'UTR or of a 3'UTR that is competent for SET recruitment may increase the surface localization of these proteins that are difficult to express.

In addition to regulating protein localization, the scaffold function of 3'UTRs was also shown to determine protein function, at least for CD47. CD47-LU interacts with SET and RAC1, which resulted in colocalization with RAC1 at the plasma membrane, RAC1 hyperactivation, lamellipodium formation, and cell migration [43]. By contrast, CD47-SU did not interact with SET or RAC1 and lacked all RAC1-mediated functions, despite also being localized to the plasma membrane like CD47-LU [43]. These findings demonstrate that proteins with identical amino acid sequences that localize to the same cellular compartment, but are generated by alternative 3'UTRs, can perform alternative functions due to the generation of alternative protein complexes. Thus, alternative 3'UTRs can diversify protein functions without changing the amino acid sequence.

It is currently unknown how widespread is the scaffold function of 3'UTRs. Many RBPs other than HuR have domains that mediate protein–protein interactions [77–81]. Thus, it is likely that HuR is not the only RBP able to recruit effector proteins to the site of translation to direct the functions of nascent proteins. It is also likely that the scaffold function of 3'UTRs is not restricted to the regulation of membrane proteins but also affects cytosolic or nuclear proteins. It can be speculated that 3'UTR-recruited effector proteins may have diverse functions. They may act as chaperones and result in alternative protein folds or they may help with the formation of multiprotein complexes by facilitating protein–protein interactions (Figure 2E). Furthermore, it can be imagined that they act as enzymes that add alternative post-translational modifications to the nascent protein, which was shown to be the case for CEBPB [82].

Finally, it can be speculated that the scaffold function may play a role in miRNA-mediated gene regulation. All human AGO genes have long 3'UTRs and generate alternative 3'UTRs [13]. miRNAs loaded into AGO are known to bind to 3'UTRs and repress gene expression [40,58]. However, miRNA effects are often subtle; therefore, it is possible that AGO may act similarly to other RBPs and recruit diverse effector proteins to nascent peptide chains to change the function of the nascent proteins.

RNA Editing and Alternative 3'UTRs Contribute to Phenotypic Diversity

RNA editing changes the sequence of the RNA without changing the DNA. Thus, mRNA editing contributes to individual variation in gene expression and results in diversification of the

transcriptome within a tissue or cell type. Interestingly, regardless of the type of RNA editing – A to I by ADAR [83], C to U by APOBEC1 [84], mRNA pseudouridilation [85,86], or m6A modifications [87,88] – all modifications show enrichment in 3'UTRs. m6A modifications were thought to be enriched surrounding stop codons [87,88]. However, a more detailed study revealed that m6A modifications mark the beginning of the terminal exon and thus may be involved in determining the end of a transcript [89]. An important feature of RNA modifications is that not all transcripts are modified, which suggests a role for RNA editing in diversification of the transcriptome.

Diversification of the transcriptome can also be achieved by the expression of alternative 3'UTRs. Several 3'-end sequencing studies showed that the number of 3'UTR isoforms generated by a gene depends on the stringency of the cut offs used to call the presence of an isoform [6,12,13]. When only robustly expressed alternative 3'UTR isoforms were considered, the fraction of human genes that generated alternative 3'UTRs was about 50% [13]. However, if minor isoforms were also included, up to 79% of mouse or human genes produced alternative 3'UTRs and generated, on average, four isoforms [12,13]. Extensive alternative 3'UTR isoform diversity was also seen in clonal populations such as cell lines or exponentially growing yeast [12,13,90]. Furthermore, mapping of 3'UTR isoform expression in single cells from mouse embryonic stem cells and neural stem cells revealed that although the developmental state globally determined isoform expression, single cells from the same state often differed in the choice of isoform [91]. Genes with higher variability in isoform expression were often moderately expressed, which suggests that many of these isoforms are expressed at less than one copy per cell. As a result, different transcriptomes may be expressed in different single cells (Figure 2F) [91]. Taking these findings together, 3'UTRs use RNA editing or alternative cleavage and polyadenylation to diversify the transcriptome and thus contribute to epigenetic gene regulation.

Concluding Remarks

3'UTR-mediated gene regulation seems to have expanded during evolution and appears to be an integral part of the functional diversity seen in higher organisms. The multifunctionality of proteins facilitated by alternative 3'UTRs may contribute to increased biological complexity. To fully understand the functional roles of 3'UTRs, it will be important to determine all the RBPs that bind to specific 3'UTRs (see Outstanding Questions). Two similar methods that identify the RBPs associated with long noncoding RNAs were recently established and identified the proteins bound to Xist [92,93]. These methods can also be applied to identify RBPs bound to specific 3'UTRs.

RBPs are highly interactive at the RNA and protein levels [80,81,94]. In addition to identifying the RBPs that bind to specific 3'UTRs, it will be important to determine the protein interaction partners of these RBPs. As illustrated above, the interaction partners of the RBPs are ultimately responsible for the different functions that are mediated by each 3'UTR. If the interaction partner is a motor protein, the mRNA will localize to different subcellular compartments, while if the interaction partner is a deadenylase or a translation initiation factor, the stability or the translation rate of the mRNA may change [48,58].

One challenge will be to dissect the different functions that can be performed by a single RBP. This is illustrated for HuR. HuR can regulate alternative polyadenylation [95,96], stabilize transcripts [97], increase translation [98], or regulate membrane protein localization [43]. It will be important to understand how a specific function of an RBP is selected to act on a single mRNA. One possibility is that post-translational modifications or protein interaction partners determine context-specific functions of RBPs [99,100].

The recent identification of approximately 800 RBPs in a cell [80,81], with most of them binding 3'UTRs, hints at a myriad of possible interactions that ultimately determine the individual fate of each mRNA. In addition to well-known RNA-binding motifs, RBPs are also strongly enriched in

Outstanding Questions

What are all the RBPs that bind to a specific 3'UTR and how does the set of bound proteins differ in various cellular contexts? Methods that were established to identify the composition and dynamics of proteins bound to long noncoding RNAs can be applied to identify the ensemble of proteins bound to specific 3'UTRs. Comparison of the 3'UTR-bound proteins across cell types will help to elucidate the contribution of 3'UTRs in the regulation of cell type-specific gene expression and function.

What are the protein interaction partners of 3'UTR-bound RBPs? Ultimately, the protein interaction partners of RBPs determine the functional outcomes of 3'UTRs. Therefore, the identification of protein interaction partners of RBPs will be necessary to address the diverse biological roles of 3'UTRs.

What determines the specific function of an RBP that acts on a single mRNA? As was illustrated for HuR, an RBP can accomplish diverse functions in gene regulation. It will be important to understand what regulates the specific function of an RBP in the context of a single gene. Post-translational modifications, protein interaction partners, and RNA granule composition have the potential to be regulators of specific RBP functions.

The scaffold function of 3'UTRs seems to diversify protein functions. Is the scaffold function of 3'UTRs also used by non-membrane proteins to alter protein function? Do all 3'UTRs mediate protein-protein interactions?

intrinsically disordered regions (IDRs) [80]. Proteins with IDRs, together with mRNAs or small RNAs, are the major components of RNA granules [101,102]. Overexpression of IDRs of known RBPs resulted in liquid phase transitions leading to the generation of RNA granules or RNA droplets because of their liquid-like behavior [101,103,104]. RNA granules can contain 100-fold higher concentrations of proteins with IDRs than the surrounding cytoplasm [105] and thus may facilitate many mRNA-based processes such as signaling, splicing, degradation, storage, and transport of mRNAs [105,106]. The purification of the mRNA components of RNA granules showed that the average 3'UTR length within granules was over fivefold longer than the average 3'UTR length found in the cytoplasm [102]. Intriguingly, the 3'UTR length of mRNAs within RNA granules is similar to the length of long 3'UTR isoforms, whereas single-UTR genes or mRNAs with short 3'UTRs seem to be present in the cytoplasm. These findings suggest that mRNAs with long 3'UTRs, which seem to be covered with RBPs, predominantly reside and carry out their functions within RNA granules. To understand the biology of 3'UTRs, it will be important to view them in the context of the bound RBPs and within diverse types of RNA granules.

Acknowledgments

The author apologizes to all colleagues whose work could not be cited because of space constraints. She is grateful to Eric Kallin, Binyamin Berkovits, Nikolaus Rajewsky, and Nicholas Proudfoot for critical comments on the manuscript. Funding for work on alternative 3'UTRs was obtained from the MSC core grant (P30 CA008748), the NCI grant U01-CA164190, the Starr Cancer Foundation, the Sidney Kimmel Cancer Foundation, and the Innovator Award of the Damon Runyon-Rachleff Cancer Foundation and the Island Outreach Foundation.

References

- Tian, B. *et al.* (2005) A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res.* 33, 201–212
- Sandberg, R. *et al.* (2008) Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* 320, 1643–1647
- Flavell, S.W. *et al.* (2008) Genome-wide analysis of MEF2 transcriptional program reveals synaptic target genes and neuronal activity-dependent polyadenylation site selection. *Neuron* 60, 1022–1038
- Ji, Z. *et al.* (2009) Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. *Proc. Natl. Acad. Sci. U.S.A.* 106, 7028–7033
- Mayr, C. and Bartel, D.P. (2009) Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* 138, 673–684
- Jan, C.H. *et al.* (2011) Formation, regulation and evolution of *Caenorhabditis elegans* 3'UTRs. *Nature* 469, 97–101
- Shepard, P.J. *et al.* (2011) Complex and dynamic landscape of RNA polyadenylation revealed by PAS-seq. *RNA* 17, 761–772
- Fu, Y. *et al.* (2011) Differential genome-wide profiling of tandem 3'UTRs among human breast cancer and normal cells by high-throughput sequencing. *Genome Res.* 21, 741–747
- Derti, A. *et al.* (2012) A quantitative atlas of polyadenylation in five mammals. *Genome Res.* 22, 1173–1183
- Lin, Y. *et al.* (2012) An in-depth map of polyadenylation sites in cancer. *Nucleic Acids Res.* 40, 8460–8471
- Elkon, R. *et al.* (2012) E2F mediates enhanced alternative polyadenylation in proliferation. *Genome Biol.* 13, R59
- Hoque, M. *et al.* (2013) Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. *Nat. Methods* 10, 133–139
- Lianoglou, S. *et al.* (2013) Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes Dev.* 27, 2380–2396
- Licatalosi, D.D. *et al.* (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 456, 464–469
- Kaida, D. *et al.* (2010) U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* 468, 664–668
- Jenal, M. *et al.* (2012) The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* 149, 538–553
- de Klerk, E. *et al.* (2012) Poly(A) binding protein nuclear 1 levels affect alternative polyadenylation. *Nucleic Acids Res.* 40, 9089–9101
- Berg, M.G. *et al.* (2012) U1 snRNP determines mRNA length and regulates isoform expression. *Cell* 150, 53–64
- Martin, G. *et al.* (2012) 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
- Bava, F.A. *et al.* (2013) CPEB1 coordinates alternative 3'-UTR formation with translational regulation. *Nature* 495, 121–125
- Lackford, B. *et al.* (2014) Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. *EMBO J.* 33, 878–889
- Masamha, C.P. *et al.* (2014) CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. *Nature* 510, 412–416
- Di Giandomartino, D.C. *et al.* (2014) RBBP6 isoforms regulate the human polyadenylation machinery and modulate expression of mRNAs with AU-rich 3' UTRs. *Genes Dev.* 28, 2248–2260
- Batra, R. *et al.* (2014) Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol. Cell* 56, 311–322
- Elkon, R. *et al.* (2013) Alternative cleavage and polyadenylation: extent, regulation and function. *Nat. Rev. Genet.* 14, 496–506
- Tian, B. and Manley, J.L. (2013) Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem. Sci.* 38, 312–320
- Bentley, D.L. (2014) Coupling mRNA processing with transcription in time and space. *Nat. Rev. Genet.* 15, 163–175
- Zheng, D. and Tian, B. (2014) RNA-binding proteins in regulation of alternative cleavage and polyadenylation. *Adv. Exp. Med. Biol.* 825, 97–127
- Han, J.D. *et al.* (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* 430, 88–93

30. Nooren, I.M. and Thornton, J.M. (2003) Diversity of protein-protein interactions. *EMBO J.* 22, 3486–3492
31. Lander, E.S. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409, 860–921
32. Hillier, L.W. *et al.* (2005) Genomics in *C. elegans*: so many genes, such a little worm. *Genome Res.* 15, 1651–1660
33. Ezkurdia, I. *et al.* (2014) Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. *Hum. Mol. Genet.* 23, 5866–5878
34. Spieth, J. *et al.* (2014) Overview of gene structure in *C. elegans*. *Wormbook* http://www.wormbook.org/chapters/www_overviewgenestructure.2/genestructure.html
35. Zhang, J. *et al.* (2008) Constraints imposed by non-functional protein-protein interactions on gene expression and proteome size. *Mol. Syst. Biol.* 4, 210
36. Levy, E.D. *et al.* (2012) Cellular crowding imposes global constraints on the chemistry and evolution of proteomes. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20461–20466
37. Jambhekar, A. and Derisi, J.L. (2007) *Cis*-acting determinants of asymmetric, cytoplasmic RNA transport. *RNA* 13, 625–642
38. An, J.J. *et al.* (2008) Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell* 134, 175–187
39. Lau, A.G. *et al.* (2010) Distinct 3'UTRs differentially regulate activity-dependent translation of brain-derived neurotrophic factor (BDNF). *Proc. Natl. Acad. Sci. U.S.A.* 107, 15945–15950
40. Bartel, D.P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215–233
41. Andreassi, C. and Riccio, A. (2009) To localize or not to localize: mRNA fate is in 3'UTR ends. *Trends Cell Biol.* 19, 465–474
42. Jambor, H. *et al.* (2015) Systematic imaging reveals features and changing localization of mRNAs in *Drosophila* development. *Elife* Published online April 2, 2015. <http://dx.doi.org/10.7554/eLife.05003>
43. Berkovits, B.D. and Mayr, C. (2015) Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. *Nature* 522, 363–367
44. Sood, P. *et al.* (2006) Cell-type-specific signatures of microRNAs on target mRNA expression. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2746–2751
45. Smibert, P. *et al.* (2012) Global patterns of tissue-specific alternative polyadenylation in *Drosophila*. *Cell Rep.* 1, 277–289
46. Ulitsky, I. *et al.* (2012) Extensive alternative polyadenylation during zebrafish development. *Genome Res.* 22, 2054–2066
47. Bertrand, E. *et al.* (1998) Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* 2, 437–445
48. Niedner, A. *et al.* (2014) Of social molecules: the interactive assembly of ASH1 mRNA-transport complexes in yeast. *RNA Biol.* 11, 998–1009
49. Ephrussi, A. *et al.* (1991) Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell* 66, 37–50
50. Greenberg, M.E. *et al.* (2009) New insights in the biology of BDNF synthesis and release: implications in CNS function. *J. Neurosci.* 29, 12764–12767
51. de Moor, C.H. *et al.* (2005) Mechanisms of translational control by the 3' UTR in development and differentiation. *Semin. Cell Dev. Biol.* 16, 49–58
52. Mazan-Mamczarz, K. *et al.* (2006) Translational repression by RNA-binding protein TIAR. *Mol. Cell Biol.* 26, 2716–2727
53. Battelli, C. *et al.* (2006) The RNA-binding protein Musashi-1 regulates neural development through the translational repression of p21WAF-1. *Mol. Cell Neurosci.* 31, 85–96
54. Pinto, P.A. *et al.* (2011) RNA polymerase II kinetics in Polo polyadenylation signal selection. *EMBO J.* 30, 2431–2444
55. Barreau, C. *et al.* (2005) AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Res.* 33, 7138–7150
56. Chen, C.Y. and Shyu, A.B. (1995) AU-rich elements: characterization and importance in mRNA degradation. *Trends Biochem. Sci.* 20, 465–470
57. Liao, B. *et al.* (2007) Competitive binding of AUF1 and TIAR to MYC mRNA controls its translation. *Nat. Struct. Mol. Biol.* 14, 511–518
58. Jonas, S. and Izaurralde, E. (2015) Towards a molecular understanding of microRNA-mediated gene silencing. *Nat. Rev. Genet.* 16, 421–433
59. Brewer, G. (1991) An A + U-rich element RNA-binding factor regulates c-myc mRNA stability *in vitro*. *Mol. Cell Biol.* 11, 2460–2466
60. Dani, C. *et al.* (1984) Extreme instability of myc mRNA in normal and transformed human cells. *Proc. Natl. Acad. Sci. U.S.A.* 81, 7046–7050
61. Schiavi, S.C. *et al.* (1992) Regulation of proto-oncogene mRNA stability. *Biochim. Biophys. Acta* 1114, 95–106
62. Stocklin, G. *et al.* (2003) A novel mechanism of tumor suppression by destabilizing AU-rich growth factor mRNA. *Oncogene* 22, 3554–3561
63. Tessier, C.R. *et al.* (2004) Mammary tumor induction in transgenic mice expressing an RNA-binding protein. *Cancer Res.* 64, 209–214
64. Vlasova, I.A. *et al.* (2008) Conserved GU-rich elements mediate mRNA decay by binding to CUG-binding protein 1. *Mol. Cell* 29, 263–270
65. Gupta, I. *et al.* (2014) Alternative polyadenylation diversifies post-transcriptional regulation by selective RNA-protein interactions. *Mol. Syst. Biol.* 10, 719
66. Spies, N. *et al.* (2013) 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. *Genome Res.* 23, 2078–2090
67. Gruber, A.R. *et al.* (2014) Global 3' UTR shortening has a limited effect on protein abundance in proliferating T cells. *Nat. Commun.* 5, 5465
68. Geisberg, J.V. *et al.* (2014) Global analysis of mRNA isoform half-lives reveals stabilizing and destabilizing elements in yeast. *Cell* 156, 812–824
69. Oikonomou, P. *et al.* (2014) Systematic identification of regulatory elements in conserved 3' UTRs of human transcripts. *Cell Rep.* 7, 281–292
70. Kristjansdottir, K. *et al.* (2015) Systematic analysis of the Hmga2 3' UTR identifies many independent regulatory sequences and a novel interaction between distal sites. *RNA* 21, 1346–1360
71. Pelechano, V. *et al.* (2013) Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature* 497, 127–131
72. Boutet, S.C. *et al.* (2012) Alternative polyadenylation mediates microRNA regulation of muscle stem cell function. *Cell Stem Cell* 10, 327–336
73. Nam, J.W. *et al.* (2014) Global analyses of the effect of different cellular contexts on microRNA targeting. *Mol. Cell* 53, 1031–1043
74. Brennan, C.M. *et al.* (2000) Protein ligands to HuR modulate its interaction with target mRNAs *in vivo*. *J. Cell Biol.* 151, 1–14
75. ten Klooster, J.P. *et al.* (2007) Rac1-induced cell migration requires membrane recruitment of the nuclear oncogene SET. *EMBO J.* 26, 336–345
76. Dunham, J.H. and Hall, R.A. (2009) Enhancement of the surface expression of G protein-coupled receptors. *Trends Biotechnol.* 27, 541–545
77. Bandziulis, R.J. *et al.* (1989) RNA-binding proteins as developmental regulators. *Genes Dev.* 3, 431–437
78. Lukong, K.E. and Richard, S. (2003) Sam68, the KH domain-containing superSTAR. *Biochim. Biophys. Acta* 1653, 73–86
79. Lunde, B.M. *et al.* (2007) RNA-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell Biol.* 8, 479–490
80. Castello, A. *et al.* (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149, 1393–1406
81. Baltz, A.G. *et al.* (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell* 46, 674–690
82. Basu, S.K. *et al.* (2011) 3'UTR elements inhibit Ras-induced C/EBP β post-translational activation and senescence in tumour cells. *EMBO J.* 30, 3714–3728

83. Peng, Z. *et al.* (2012) Comprehensive analysis of RNA-seq data reveals extensive RNA editing in a human transcriptome. *Nat. Biotechnol.* 30, 253–260
84. Rosenberg, B.R. *et al.* (2011) Transcriptome-wide sequencing reveals numerous APOBEC1 mRNA-editing targets in transcript 3' UTRs. *Nat. Struct. Mol. Biol.* 18, 230–236
85. Schwartz, S. *et al.* (2014) Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* 159, 148–162
86. Carlile, T.M. *et al.* (2014) Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* 515, 143–146
87. Dominissini, D. *et al.* (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206
88. Meyer, K.D. *et al.* (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149, 1635–1646
89. Ke, S. *et al.* (2015) A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev.* 29, 2037–2053
90. Pelechano, V. *et al.* (2012) Genome-wide polyadenylation site mapping. *Methods Enzymol.* 513, 271–296
91. Velten, L. *et al.* (2015) Single-cell polyadenylation site mapping reveals 3' isoform choice variability. *Mol. Syst. Biol.* 11, 812
92. Chu, C. *et al.* (2015) Systematic discovery of Xist RNA binding proteins. *Cell* 161, 404–416
93. McHugh, C.A. *et al.* (2015) The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* 521, 232–236
94. Kim, P.M. *et al.* (2008) The role of disorder in interaction networks: a structural analysis. *Mol. Syst. Biol.* 4, 179
95. Hilgers, V. *et al.* (2012) ELAV mediates 3' UTR extension in the *Drosophila* nervous system. *Genes Dev.* 26, 2259–2264
96. Oktaba, K. *et al.* (2015) ELAV links paused Pol II to alternative polyadenylation in the *Drosophila* nervous system. *Mol. Cell* 57, 341–348
97. Fan, X.C. and Steitz, J.A. (1998) Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the *in vivo* stability of ARE-containing mRNAs. *EMBO J.* 17, 3448–3460
98. Mazan-Mamczarz, K. *et al.* (2003) RNA-binding protein HuR enhances p53 translation in response to ultraviolet light irradiation. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8354–8359
99. Liu, Q. and Dreyfuss, G. (1995) *In vivo* and *in vitro* arginine methylation of RNA-binding proteins. *Mol. Cell. Biol.* 15, 2800–2808
100. Siomi, H. and Dreyfuss, G. (1997) RNA-binding proteins as regulators of gene expression. *Curr. Opin. Genet. Dev.* 7, 345–353
101. Kato, M. *et al.* (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* 149, 753–767
102. Han, T.W. *et al.* (2012) Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. *Cell* 149, 768–779
103. Kim, H.J. *et al.* (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495, 467–473
104. Nott, T.J. *et al.* (2015) Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* 57, 936–947
105. Li, P. *et al.* (2012) Phase transitions in the assembly of multivalent signalling proteins. *Nature* 483, 336–340
106. Weber, S.C. and Brangwynne, C.P. (2012) Getting RNA and protein in phase. *Cell* 149, 1188–1191