

REVIEW ARTICLE

# Alternative polyadenylation of mRNA and its role in cancer



Fuwen Yuan <sup>a</sup>, William Hankey <sup>a</sup>, Eric J. Wagner <sup>b</sup>, Wei Li <sup>c</sup>,  
Qianben Wang <sup>a,d,\*</sup>

<sup>a</sup> Department of Pathology, Duke University School of Medicine, Durham, NC, 27710, USA

<sup>b</sup> Department of Biochemistry and Molecular Biology, University of Texas Medical Branch at Galveston, Galveston, TX, 77555, USA

<sup>c</sup> Department of Biological Chemistry, University of California, Irvine, CA, 92697, USA

<sup>d</sup> Duke Cancer Institute Center for Prostate and Urologic Cancers, Duke University School of Medicine, Durham, NC, 27710, USA

Received 15 October 2019; accepted 18 October 2019

Available online 25 October 2019

## KEYWORDS

3' untranslated region;  
Alternative polyadenylation;  
Cancer;  
Gene regulation;  
Polyadenylation signals

**Abstract** Alternative polyadenylation (APA) is a molecular process that generates diversity at the 3' end of RNA polymerase II transcripts from over 60% of human genes. APA is derived from the existence of multiple polyadenylation signals (PAS) within the same transcript, and results in the differential inclusion of sequence information at the 3' end. While APA can occur between two PASs allowing for generation of transcripts with distinct coding potential from a single gene, most APA occurs within the untranslated region (3'UTR) and changes the length and content of these non-coding sequences. APA within the 3'UTR can have tremendous impact on its regulatory potential of the mRNA through a variety of mechanisms, and indeed this layer of gene expression regulation has profound impact on processes vital to cell growth and development. Recent studies have particularly highlighted the importance of APA dysregulation in cancer onset and progression. Here, we review the current knowledge of APA and its impacts on mRNA stability, translation, localization and protein localization. We also discuss the implications of APA dysregulation in cancer research and therapy.

Copyright © 2019, Chongqing Medical University. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

\* Corresponding author. Department of Pathology, Duke University School of Medicine, Durham, NC, 27710, USA. Fax: +1 9193857554.  
E-mail address: [qianben.wang@duke.edu](mailto:qianben.wang@duke.edu) (Q. Wang).  
Peer review under responsibility of Chongqing Medical University.

## Introduction

It is well-established that mRNA conveys genetic information from DNA to the protein level.<sup>1</sup> Newly synthesized precursor mRNA (pre-mRNA) undergoes multiple type of RNA processing events to lead to production of mature mRNA, which is ultimately translocated to the ribosome to act as a protein translation template. While alternative splicing has been known to generate diversity within transcripts, the process of transcriptional termination was long thought to be a constitutive event. Termination is an obligatory RNA processing occurrence and includes a cleavage step in which sequence is enzymatically removed from the 3' end of the transcript, followed by the addition of a series of adenosine monophosphate units to form a poly(A) tail that plays a role in the transcript's stability, nuclear export, and translation. Work over the past two decades have enabled the realization that the majority of human pre-mRNAs contain multiple alternative PAS sequences. This suggests that a single gene is capable of utilizing multiple cleavage and polyadenylation signals, giving rise to two or more mRNA transcripts with distinct 3' ends.<sup>2</sup> Therefore APA is critical for controlling mRNA stability, localization, translation, protein coding and protein localization, and emerging evidence suggests that APA also plays key roles in the regulation of gene expression and gene function.<sup>3–5</sup> Similar to splicing, APA is regulated by *cis*-regulatory elements within the pre-mRNA that are recognized by *trans*-regulators. These elements include: the poly-A signal itself, typically AAUAAA or its variants (termed PAS), which is recognized by cleavage and polyadenylation specificity factors (CPSFs),<sup>6,7</sup> auxiliary UGUA sequence elements recognized by members of the mammalian cleavage factor I (CFIm),<sup>8</sup> and GU-/U-rich downstream sequences bound by cleavage stimulation factors (CSTFs).<sup>9,10</sup> While the full repertoire of proteins within the cleavage and polyadenylation (CPA) complex may exceed 80 members, the typical core APA regulation complex consists of CPSFs, CSTFs, CFIm and the mammalian cleavage factor complexes II (CFIIm).<sup>11</sup> In addition to these complexes, several single factors are also required for the cleavage and polyadenylation process including poly(A) polymerase (PAP) and the scaffolding proteins poly(A)-binding protein (PAB) and Symplekin.<sup>11–13</sup>

Global PAS analyses using a variety of next-generation sequencing (NGS) approaches estimate that over one-third of mouse genes and two-thirds of human genes utilize more than one PAS, indicating that the generation of different APA transcripts from a single gene in different cellular contexts is not just a potential event but is a widespread phenomenon.<sup>14–17</sup> APA has been found to occur naturally during a variety of cellular processes including during development,<sup>18–20</sup> in response to stress response,<sup>21,22</sup> changes in cell growth status,<sup>23,24</sup> in different tissues<sup>25</sup> and pathologically in normal versus cancer cells/tissues.<sup>26,27</sup> As compelling examples, well-differentiated cells like neurons have a general tendency to use PASs distally located relative to stop codons, resulting in the expression of transcripts with longer 3'UTR.<sup>28,29</sup> On the other hand, faster growing cells tend to use the proximal PAS creating shorter 3'UTR resulting in potential higher expression

levels.<sup>26,30–32</sup> From a physiological perspective, Sandberg et al demonstrated that activated T cells with genes bearing shorter 3'UTR exhibit growth advantages relative to their counterparts with longer 3'UTR, and 3'UTR shortening was similarly observed after human monocytes or B cells were stimulated.<sup>23</sup> In addition, proximal 3'UTR PAS usage is increased in multiple cancer cell lines and tissues and promotes cancer proliferation.<sup>26,33</sup> These findings paint a picture that highlights the importance of APA in areas of cellular physiology ranging from stress response to cellular differentiation, and most importantly cancer progression. This review focuses on the general concept of APA and its roles in gene regulation as it pertains to tumorigenesis.

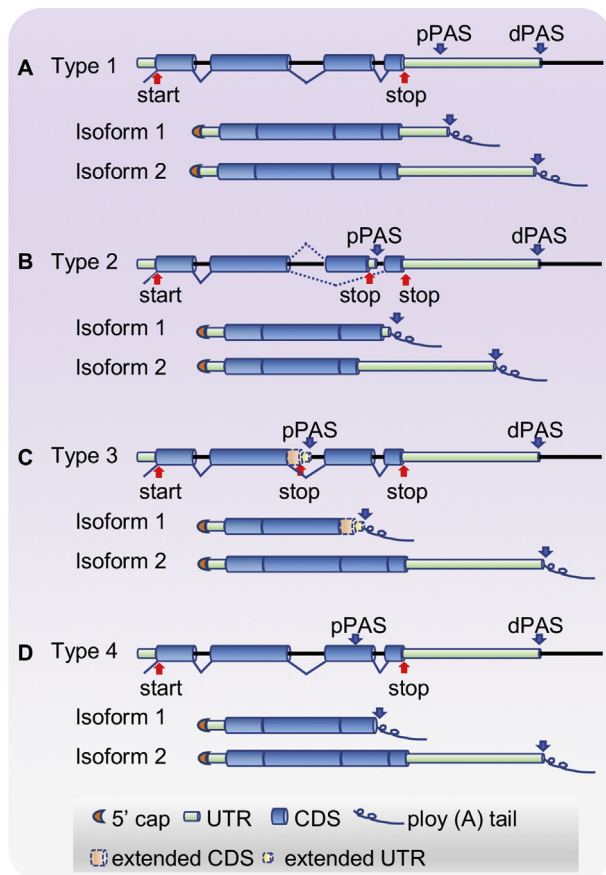
## Mechanistic regulation of APA

### Different types of APA

The distinct 3'terminal region of each APA-generated isoform is determined by the location of the PAS and, as such, APA can be broken down into four subclasses.<sup>2,34</sup> 'Tandem 3'UTR APA' is the simplest and most common type, involving cleavage within the 3'UTR resulting in altered 3'UTR length but does not result in any change to coding potential. 'Splicing-APA' or alternative terminal exon polyadenylation is the second most widely-occurring type that leads to transcripts with entirely distinct 3'UTR sequences as well as different C-terminal amino acids of the encoded protein.<sup>35,36</sup> The last two types, which are less common, are 'intronic APA' and 'internal exon APA' which occur upstream of the annotated stop codon and can lead to production of truncated proteins (Fig. 1). Importantly, as both intronic APA and internal exon APA can be subject to mRNA decay pathways such as nonsense-mediated decay (NMD) or non-stop decay, estimates of their frequency may be low. The alternative splicing machinery may also contribute APA further increasing transcript diversity generated from a single gene.<sup>37</sup> Regardless of which APA type occurs, there is the potential to alter the presence of specific *cis*-regulatory elements located in the distal 3'UTR, which are known to determine mRNA stability, efficiency of translation, nuclear localization, and even the localization of the encoded protein.<sup>26,38–41</sup> One remarkable feature of 3'UTR APA is that it has been shown to change globally during cancer progression and in different cellular and tissue models of cancer.<sup>26,27,33,42</sup> This seems surprising given the complexity of outcomes that altering APA on global scale would generate but clearly provides growth advantages to tumors. Regardless of the root cause, it is clear that understanding the mechanistic regulation of 3'UTR APA and identifying the key factors responsible for 3'UTR APA changes during tumorigenesis and cancer progression are of great importance for the discovery of potential alternative tumor therapies.

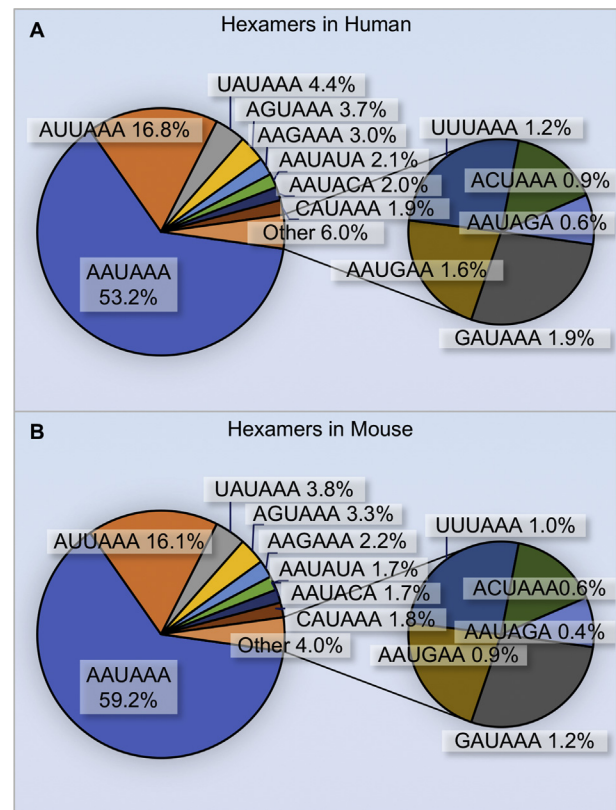
### Cis elements required for APA

Alignment of human and mouse cDNA/EST sequences near their 3' end allows for readily identifiable specific motifs potentially required for correct PAS recognition and



**Figure 1** Four types of alternative polyadenylation events. (A) Tandem 3'UTR APA occurs within transcripts that possess two or more cleavage/poly(A) sites within their 3' UTR, resulting in 3'UTR length differences between APA isoforms that code for identical gene products. (B) Alternative terminal exon APA is caused by alternative splicing, which results in internal exon skipping and use of a distinct poly(A) site contained within the new terminal exon. (C) Intronic APA involves the use of cryptic alternative poly(A) sites found within introns. (D) Internal exon APA features Poly(A) site usage within upstream exons, resulting in a truncated mRNA isoform lacking both a stop codon and a 3'UTR.

processing (Fig. 2). This approach identifies that the most prominent elements in both human and mouse transcripts are PAS hexamers, which appear as AAUAAA in approximately 50% of human and 60% of mouse transcripts.<sup>6,7</sup> Nine additional PAS variants have been identified and collectively appear in ~30% of PAS sites. These hexamer variants are functionally weaker at directing polyadenylation compared to the consensus AAUAAA hexamers and may even be inactive *in vitro*.<sup>43</sup> Recently, several structural studies have uncovered the molecular basis for the AAUAAA preference and these structures underscore how just single variants in this motif can dramatically reduce affinity.<sup>44–46</sup> Despite the relative intolerance of AAUAAA base alterations, variants such as AAGAAA are very common, especially for proximal/upstream APA sites. This suggests that other elements and protein factors may impact recognition



**Figure 2** Top detected PAS hexamers in human and mouse. (A and B) Human (A) and mouse (B) genomic sequences located between  $-40$  and  $-1$  nucleotides upstream of poly(A) sites were used to detect hexamers that might function as polyadenylation signals.

of a suboptimal PAS. Indeed, several other *cis* elements located upstream and downstream of the PAS are also necessary for correct and efficient PAS recognition, cleavage and polyadenylation. Most commonly, UGUA elements are usually found within 40 nt upstream of AAUAAA and function to recruit and bind subunits of the CFI complex.<sup>47</sup> Moreover, the cleavage site itself is most commonly a CA dinucleotide that is often found about 15–30 bp downstream of PAS and is recognized and cut by CPSF73.<sup>48–52</sup> Finally, the downstream sequence element (DSE) is a U-rich or GU-rich motif that appears frequently and is recognized by subunits of the CSTF complex.<sup>53</sup> All of these sequence motifs are usually contained within a region of 100 nt, as downstream U/GU-rich elements and upstream UGUA elements are located near the cleavage site and usually within 40 nt of the PAS.<sup>47,54</sup> Collectively, the overall adherence to consensus and number of UGUA motifs is thought to establish how efficiently a given PAS is utilized.

### Trans-regulators involved in APA

More than 80 APA-related factors have been identified in the human pre-mRNA 3' processing complex.<sup>11</sup> Polyadenylation machinery in metazoans canonically consists of ~20 core proteins<sup>5</sup> that make up the CPSF, CSTF, CFI and CFII complexes (Fig. 3).

### CPSF

The CPSF complex contains six subunits and includes: the CPSF 160 kDa subunit (CPSF160), CPSF100, CPSF73, CPSF30, WDR33 and FIP1.<sup>11,13,55,56</sup> Overall, this complex is functionally separated into two submodules: 1) components that recognize the AAUAAA hexamer within the mammalian PAS and mainly act as an anchor for other 3' processing factors; 2) components that function in that actual catalysis of RNA cleavage. More specifically, WDR33 and CPSF30 recognize and directly bind to the AAUAAA hexamer<sup>57–59</sup> while CPSF160 coordinates the assembly of other CPSF components to facilitate AAUAAA binding.<sup>44,45,56</sup> FIP1 binds to U-rich sequence upstream of the AAUAAA hexamer *in vitro* and *in vivo* via its Arg-rich C-terminal domain to further modulate PAS recognition. Consistent with a potential regulatory function, FIP1 was reported to regulate APA in embryonic stem cells (ESCs) and is important for ESC self-renewal.<sup>60</sup> FIP1 and CPSF160 also function to directly facilitate the accumulation of poly(A) polymerase (PAP) at the mRNA 3' processing site.<sup>12,13</sup> The cleavage reaction itself is carried out by CPSF100 and CPSF73, which form an obligate heterodimer with the endonuclease activity being housed within  $\beta$ -lactamase/ $\beta$ -CASP domain of CPSF73.

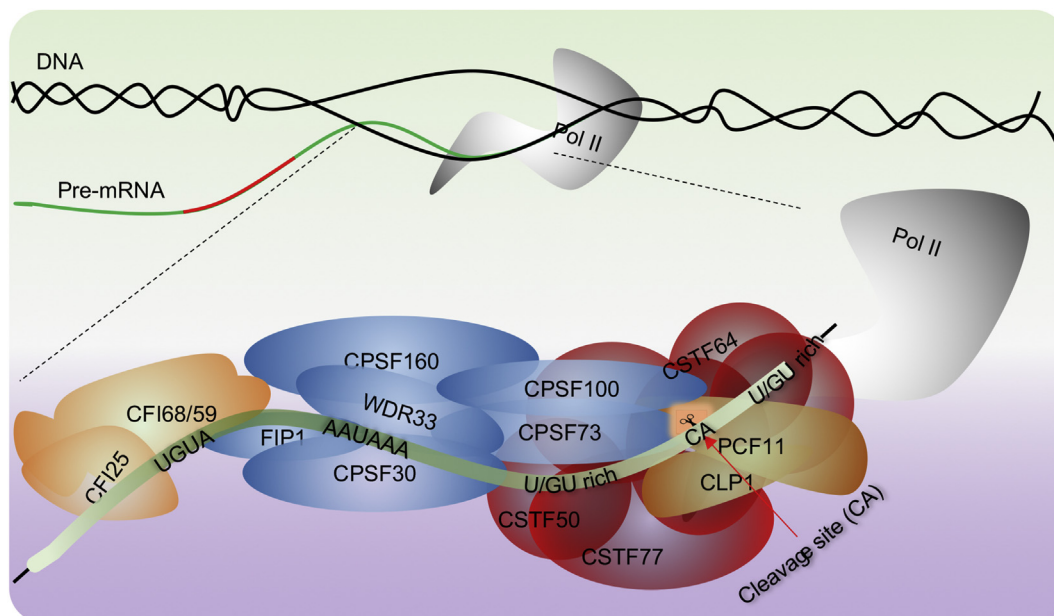
### CSTF

The CSTF complex consists of three subunits: CSTF1 (also named CSTF50), either CSTF2 (also named CSTF64) or its paralog CSTF64 $\tau$  (or another isoform  $\beta$ CSTF64 that is usually found specifically in the central nervous system<sup>61</sup>), and CSTF3 (also named CSTF77). CSTF64 or CSTF64 $\tau$  directly binds to the downstream U/GU rich element and acts as a scaffold to recruit CSTF77 and CSTF50, each of which is

thought to self-associate to form a homodimer.<sup>62–64</sup> The function of the CSTF complex is to enhance recognition of the upstream PAS by the CPSF but the molecular basis of this potential stimulation is currently unknown. Consistent with this role, modulation of the strength of the downstream element or CSTF subunit expression can have significant impact on PAS selection. Consistently, while knockdown of either CSTF64 or CSTF64 $\tau$  alone has little effect on global PAS selection, double knockdown of both CSTF64 and CSTF64 $\tau$  leads to significant APA changes.<sup>65</sup>

### CFIm and CFII complexes

The process of cleavage and polyadenylation typically requires two auxiliary complexes, termed CFIm and CFII. The better-characterized CFIm complex consists of the CFIm 25 kDa subunit (CFIm25) that forms homodimers that further heterodimerize with two copies of either CFIm68 or CFIm59 and is thought to stimulate cleavage reactions both *in vitro* and *in vivo*.<sup>8,66–68</sup> The CFIm complex can directly bind to the UGUA motifs<sup>68</sup> located nearby the PAS, through specific binding of its CFIm25. Both CFIm68 and CFIm59 also contain the ability to bind pre-mRNA but are thought to be non-specific.<sup>69</sup> The CFIm complex plays an important role in APA selection, especially for the selection preference of distal polyadenylation sites.<sup>33,47,66,70,71</sup> CFIm25 contains a critical Nudix domain that has been shown in other Nudix proteins to contain pyrophospho-hydrolase activity, but evolution has mutated/inactivated key catalytic residues rendering CFIm25 inert. Rather, this Nudix domain is responsible for interactions with the UGUA motif<sup>68</sup> and may ultimately be critical for regulating PAS selection.<sup>68,72</sup> Both CFIm59 and CFIm68 exhibit domain structures similar to



**Figure 3** *Cis* elements and core factors involved in cleavage and polyadenylation. The PAS is recognized by the CPSF complex and is directly bound by CPSF30 and WDR33. The cleavage site, usually located ~15–30 nt downstream of the PAS, can be sheared by the endonuclease CPSF73 with the help of other regulators. U/GU-rich elements are often found within 100 nt downstream of PAS hexamers, and can be recognized by the CSTF complex. The CFIm complex can specifically recognize the UGUA-rich element located within 40 nt upstream of the PAS. Several additional proteins not shown here are also required for APA including PAP, PABP1, RBBP6, symplekin and the Pol II CTD.

those of SR proteins, a family of well-known splicing regulators.<sup>73</sup> Recent work has shown that the CFIm complex functions early in the APA process and helps to stabilize CPSF binding to the PAS.<sup>69,74</sup> Given the timing of this function, cleavage factor complexes are believed to play critical roles in PAS selection. Indeed, reduction in the expression of CFIm25 or CFIm68 (but not CFIm59) leads to global 3'UTR shortening, supporting the role of the CFIm complex in determining the terminal exon definition.<sup>70,75</sup> As an example underscoring this property, glioblastoma cells exhibit decreased expression of CFIm25 accompanied by a global increase in proximal PAS usage and enhanced tumorigenicity.<sup>33,42</sup> The CFIm complex is the least characterized among 3' end processing subcomplexes, consisting of the polyadenylation factor CLP1 (also known as hClp1) and PCF11.<sup>76</sup> PCF11 preferentially binds to G-rich sequence elements of RNA with high affinity, and is necessary for the cleavage stimulation activity of CFIm.<sup>76</sup> The CFIm complex differs from the CFIm complex in that it tends to promote proximal PAS usage, with knockdown of PCF11 leading to upregulation of distal PAS isoforms.<sup>35,77</sup>

### Auxiliary APA factors

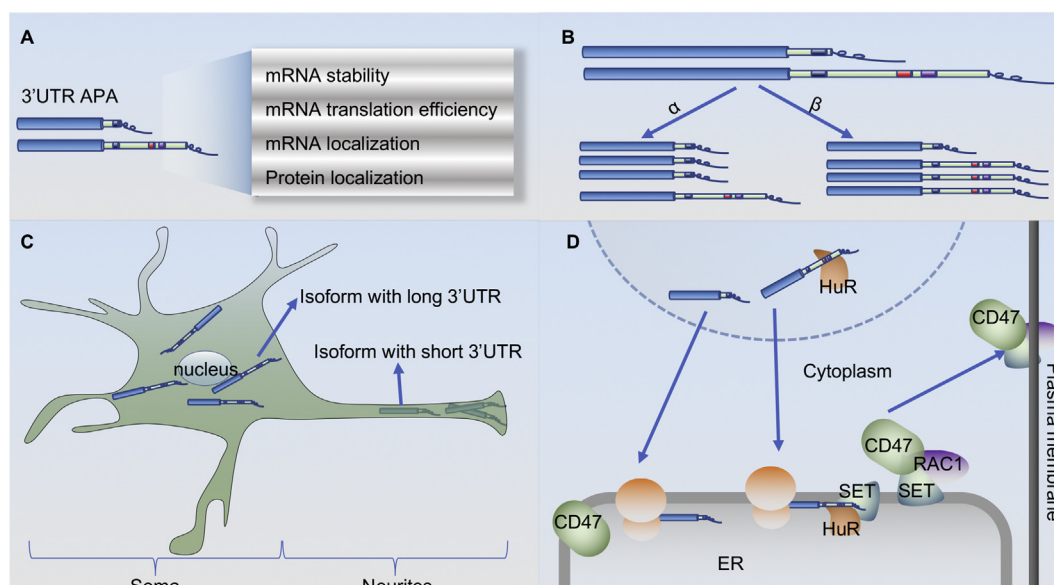
In addition to the core complexes described above, several other RNA binding proteins and other factors can regulate APA. These include the poly A polymerase (PAP) complex (which consists of PAP $\alpha$  and PAP $\gamma$ ), retinoblastoma-binding protein 6 (RBBP6), nuclear polyadenylate-binding protein 1 (PABPN1), Symplekin, heterogeneous nuclear RNP H (HNRNPH), and the phosphorylation status of the regulatory carboxy-terminal domain (CTD) of RPB1, the large subunit of RNA polymerase II (Pol II).<sup>3,11</sup> Within this group, PAP is required for efficient cleavage and is recruited by FIP1 and CPSF160 to auxiliary upstream elements.<sup>13</sup> The binding of PAP with U1 snRNP, a core component of the spliceosome,<sup>78</sup> results in inhibition of polyadenylation as U1 snRNP can negatively regulate PAP polyadenylation activity.<sup>79–81</sup> PABPN1 is a major poly(A) tail binding protein in the cytoplasm that shuttles back and forth to the nucleus and is important for synthesis of the appropriately sized poly(A) tail.<sup>82,83</sup> PABPN1 also promotes distal PAS usage, as it inhibits mRNA proximal PAS cleavage *in vitro*, and its ablation induces global 3'UTR shortening.<sup>84</sup> RBBP6 was first identified as a binding protein for p53 and Rb,<sup>85,86</sup> and more recently has been implicated in APA regulation.<sup>87</sup> This role appears to be mediated by the N-terminal region of RBBP6, which interacts with the CSTF complex and contains a DWNN/ubiquitin-like domain, a zinc knuckle and a RING finger. The zinc knuckle and RING finger domains function in RNA binding, and the N-terminal region is sufficient for stimulating 3' processing activity *in vitro*. Knockdown of RBBP6 in mammalian cells leads to broad 3'UTR lengthening and preferential inhibition of usage of PAS containing AU-rich elements within their 3'UTRs.<sup>87</sup> PAS usage is also determined in part by Pol II, with distinct PAS usage patterns resulting from alterations in Pol II elongation rate within the PAS region.<sup>51,88–92</sup> Mutations in the transcription elongation factors TFIIS and Spt5 or in the Pol II subunit Rpb2 also promote proximal PAS usage in yeast.<sup>93</sup> Finally, proteins such as Symplekin and the Pol II CTD do not bind RNA but function as scaffolds for 3' processing.<sup>11,94</sup>

## Molecular effects of APA

The most straightforward result of APA occurs when the alternative PAS is located within the protein coding region (Types 2–4 in Fig. 1) so as to alter the amino acid sequence generating a different protein isoform.<sup>35,36</sup> For instance, Lee et al found that primary chronic lymphocytic leukemia (CLL) cells exhibit widespread upregulation of truncated mRNAs and proteins resulting from intronic polyadenylation.<sup>30</sup> Aside from these less common instances, predicting the impact of 3'UTR APA (Type 1 in Fig. 1) is less intuitive as coding sequences are not altered, yet lengthening or shortening the 3'UTR affects the abundant *cis* elements that can be targeted by miRNA or RNA binding proteins (RBPs).<sup>23</sup> Therefore, the ultimate output of 3'UTR APA depends on the cellular milieu in terms of whether miRNA or RBPs are expressed in a given cell type (Fig. 4). Here we will discuss the effects of 3'UTR APA in detail.

### APA impacts mRNA stability

The 3'UTR contains *cis* elements specifically recognized by RBPs, miRNAs and lncRNAs. miRNAs are small RNAs (~22 nt) that modulate the stability and/or translation of complementary target mRNAs, and were once hypothesized to be the main factors governing the differential mRNA stability of distinct mRNA isoforms produced by APA. This assumption is supported by the fact that most miRNA target sites are located within 3'UTRs and over 50% of conserved miRNA binding sites in mRNA transcripts are collectively located within 3'UTR regions impacted by APA.<sup>23,29</sup> However, despite numerous studies that have observed that 3'UTR shortening results in a global increase in mRNA half-life, especially in cancer-derived cells and tissues,<sup>26,27</sup> Nam et al found that miRNAs may influence only 10% of APA isoform expression differences between HeLa, HEK293 and Huh7 cells.<sup>95</sup> A more recent study has further demonstrated that although 3'UTR APA results in the loss of *cis* elements from the 3'UTR, only small fraction of shortened mRNA exhibit expression differences.<sup>96</sup> One possible explanation is related to the observation that miRNAs bind more efficiently to target sites located in proximity to the location of the poly(A) tail. As a result, proximal APA usage resulting in 3'UTR shortening not only causes loss of miRNA binding elements but also increases the binding efficiency of some miRNAs that bind upstream of the proximal PAS. For example, 3'UTR shortening on antiproliferative and pro-differentiation genes during cell proliferation has been shown to enhance the targeting efficiency of miRNAs that bind upstream of the proximal PAS, decreasing mRNA stability and promoting cell proliferation.<sup>39</sup> Taken together, these findings indicate that the impact of miRNA on the stability of different transcripts generated from 3'UTR APA can vary depending on the transcript and cellular context. 3'UTR shortening on the one hand may result in loss of miRNA binding sites and increased stability of the transcript, or it may enhance the binding efficiency of miRNAs targeting sites upstream of the proximal PAS and promote 3'UTR shortened mRNA degradation. Thus, the ultimate effects of APA must be considered on a case-by-case format.



**Figure 4** Impact of 3'UTR APA on its target mRNA and protein. (A) Summary of potential consequences of 3'UTR alternative cleavage and polyadenylation on the affected mRNA and its encoded protein. (B) Long 3'UTRs contain more *cis* elements than short 3'UTRs, influencing both the stability and translation rate of the isoforms. The  $\alpha$  pathway depicts the effect of *cis* elements such as miRNA binding sites in the long 3'UTR, promoting degradation of the lengthened isoform. The  $\beta$  pathway depicts the effect of 3'UTR shortening in increasing the miRNA binding efficiency in the middle region of the 3'UTR, thus decreasing the stability of the shortened isoform. In addition, some *cis* elements in the 3'UTR contribute to translation efficiency differences between transcripts bearing a longer or shorter 3'UTR. (C) Transcripts that undergo APA at the distal PAS remain in the soma of neurons, whereas those that use proximal PAS preferentially localize to neurites. (D) *Cis*-regulatory elements in 3'UTRs can direct the cellular localization of the encoded protein by altering interactions with RNA-binding proteins. The 3'UTR of the long isoform of CD47 can be bound by HuR and recruits SET to the translation site, facilitating the nascent CD47 protein to join active RAC1 and translocate to the plasma membrane.

In addition to miRNA binding elements, 3'UTRs harbor mRNA destabilization elements that can be recognized by RBPs. These include GU-rich elements, AU-rich elements and PUF protein-binding elements.<sup>97,98</sup> The inclusion or exclusion of these elements represents yet an additional mechanism by which 3'UTR APA can affect mRNA stability. In mammalian cells, transcripts with longer 3'UTRs containing predominant destabilizing elements are generally degraded more quickly than those with shorter 3'UTRs.<sup>99–101</sup> However, the impact on mRNA stability of 3'UTR elements recognized by RBPs is far more complicated than previously thought. Other studies have found stabilizing elements, such as HuR binding element, specifically in longer 3'UTRs, which can substantially affect mRNA decay.<sup>100,102–104</sup>

In summary, the impact of 3'UTR APA on mRNA stability is highly heterogeneous and dependent on 3'UTR sequence contents and cellular conditions. This is consistent with transcriptome-wide studies that have found that although 3'UTR APA influences the decay rates of about one-third of mRNA isoforms, the isoforms with shorter 3'UTRs are only slightly more stable on average than those with longer 3'UTRs.<sup>100,102</sup> Despite this unpredictability, it is clear that 3'UTR APA produces multiple mRNA isoforms that decay at different rates, highlighting the importance of 3' UTR APA in regulating mRNA stability.

### APA on mRNA translation

In addition to containing sequences that determine mRNA stability, the 3'UTR also includes elements that interact with factors regulating mRNA translation. Indeed, studies analyzing the effects of APA on mRNA translation have reported that longer isoforms are often associated in polysome fractions and with higher translation efficiency.<sup>41,100</sup> In one example, unstimulated hippocampal neurons predominantly contain a brain-derived neurotrophic factor (BDNF) isoform with a shorter 3'UTR, while a longer, more efficiently translated 3'UTR isoform is exclusively present following neuronal activation or differentiation.<sup>41,105,106</sup> The difference in translation efficiency enables neuronal activation/differentiation to trigger a rapid accumulation of BDNF protein levels. Similarly, the *Drosophila* transcript *polo*, which encodes Polo-like kinase,<sup>90,107</sup> was demonstrated to exhibit higher translation efficiency from its longer 3'UTR isoform than for the shorter 3'UTR isoform. Loss of the distal PAS of *polo* is lethal for flies during development, while proximal PAS deletion has no phenotypic effect. These studies indicate that the higher translation rates of mRNA isoforms with longer 3'UTR are required in some cases for large-scale cell proliferation and differentiation.<sup>90</sup> Transcripts with long 3'UTR, however, do not universally exhibit a higher translation efficiency than their short 3'UTR counterparts. In fact, a recent study has

reported that in some cell lines, distal PAS usage resulting in lengthened 3'UTR decreases polysome occupancy and represses mRNA translation in comparison with shortened 3'UTR isoform.<sup>108</sup> Therefore, the impact of 3'UTR APA on mRNA translation may vary in a gene- and 3'UTR sequence-dependent manner. Further work is required to delineate how various *cis* elements affect the translation efficiency of APA isoforms in different conditions and different cell types.

### mRNA nuclear export and localization

Interactions between 3'UTR-associated RBPs and motor proteins can facilitate transcript mobilization to different subcellular compartments.<sup>109–112</sup> Recent studies have reported that ~10% of 3'UTR APA isoforms exhibit distribution differences between the nucleus and cytoplasm, with long 3'UTR isoforms generally enriched in the nucleus relative to the cytoplasm.<sup>113,114</sup> However, it is difficult to assess whether 3'UTR length difference actively contributes to nuclear accumulation or simply mediates cytoplasmic degradation of the long 3'UTR isoform. 3'UTR APA also leads to localization differences within the cytoplasm, as long 3'UTRs facilitate mRNA localization to the endoplasmic reticulum (ER) and accelerate the expression of membrane proteins.<sup>115,116</sup> In neuronal cells, symmetrically localized mRNAs can function to sequester proteins, establish cell polarity and direct asymmetric cell division.<sup>117,118</sup> As an example, in neurons, the short 3'UTR isoform encoding BDNF is restricted to and translated within the soma, while the long 3'UTR isoform is required for appropriate localization of BDNF to the dendrites. The importance of this mechanism is seen in mice lacking the long the 3'UTR isoform, which present with altered dendritic spine morphology and decreased dendritic synapse plasticity.<sup>109</sup> Similar isoform-specific distribution patterns between the axon and the cell body are observed for distinct isoforms encoding RAN<sup>119</sup> and inositol monophosphates 1.<sup>120</sup> Although the importance of 3'UTR APA in mRNA localization and downstream biological processes has been shown in these examples, further investigations are needed to delineate the underlying mechanisms. These future studies must identify the *cis* elements unique to long 3'UTR isoforms and the *trans* factors that bind to them to mediate mRNA localization differences. Since 3'UTR length simultaneously impacts mRNA stability, future studies are also necessary to clarify whether *cis* elements involved in mRNA stability also contribute to the apparent enrichment of different APA isoforms in various subcellular compartments.

### Protein localization

Surprisingly, recent studies have reported that 3'UTR APA can impact the localization of proteins generated from mRNA transcripts of different 3'UTR length, independently of mRNA localization.<sup>38,121</sup> A well-established example is the impact of 3'UTR APA on CD47 protein localization. *CD47* isoforms with short and long 3'UTR encode the same protein. The protein encoded by the shorter isoform protein localizes to the ER, while the protein encoded by the longer

isoform localizes to the plasma membrane. Mechanistic studies have revealed that the long 3'UTR can bind to the RBP human antigen R (HuR), which then recruits the effector protein SET to site of translation where it is transferred from the mRNA to the nascent CD47 protein.<sup>38</sup> SET then binds and activates RAC1 to trigger translocation of the SET/CD47 complex to the plasma membrane. TIS11B, an AU-rich binding protein, is required for 3'UTR-mediated CD47 cell surface localization and for the interaction of SET with the long 3'UTR *CD47* isoform.<sup>122</sup> While plasma membrane-associated CD47 generated from the long 3'UTR isoform mainly functions to protect cells from phagocytosis by macrophages, CD47 translated from the short 3'UTR isoform functions in the regulation of apoptosis.<sup>38</sup> Other proteins generated from alternative 3'UTR isoforms, such as CD44, TNF receptor superfamily member 13C (TNFRSF13C) and  $\alpha 1$  integrin (ITGA1), exhibit similar protein distribution differences.<sup>38</sup> Future studies may be needed to investigate the extent to which 3'UTR APA regulates protein localization in a global manner, leading to distinct functions of the same proteins localized at different places in the cell.

## APA modulation in cancer

### Widespread APA in tumorigenesis

Next-generation sequencing technologies have facilitated global mRNA profiling of a variety of human disease states and their corresponding normal tissues.<sup>123–125</sup> From the subset of those studies that permit evaluation of PAS usage, an increasing body of evidence has suggested that APA is an important regulatory mechanism for the activation of oncogenes. First, in proliferating cells, the upregulation of genes related to cell growth is accompanied by 3'UTR shortening.<sup>23,26,27,30–33,126–128</sup> Second, studies have observed preferential global truncation of 3' UTR isoforms in different cancer cell lines, indicating a strong association of APA with cancers.<sup>26</sup> Third, systems analysis of RNA-seq data from 358 sets of paired tumor and adjacent non-tumor tissues across seven tumor types from the Cancer Genome Atlas (TCGA) database has identified ~1350 genes with recurrent APA changes, with around ~90% occurring in the 3'UTR region and resulting in 3'UTR shortening. Importantly, selected APA events add strong prognostic power beyond regular clinical and molecular parameters, indicating that these APA events may serve as novel prognostic biomarkers.<sup>27,127</sup> Two recent studies of triple-negative breast cancer found that the majority of 3'UTR APA in breast cancer tissues results in 3'UTR shortening relative to normal breast tissues. Most of these genes with shortened 3'UTR in patient tissues are proliferation-related transcripts correlated with recurrence-free survival times, supporting the notion of APA-based proto-oncogene activation.<sup>129,130</sup> These data are complemented by evidence from mouse model studies supporting a critical role for APA in cancer. A study from a mouse model of B-cell leukemia/lymphoma has revealed that global 3'UTR length changes in cancer can serve as biomarkers to distinguish between similar tumor subtypes with different survival characteristics.<sup>31</sup> In addition to the direct impact of individual genes with 3'UTR APA on cancer, the diversity of competing-endogenous RNA

(ceRNA) networks generated from 3'UTR APA may also contribute to tumorigenesis by competitively altering the availability of shared miRNAs targeting them. A recent study has demonstrated that 3'UTR shortening can alter the distribution of ceRNA networks and repress tumor-suppressor genes such as *PHF6* and *LARP1*.<sup>131</sup>

### APA factors involved in cancer

Although global 3'UTR APA occurs in different cancer cells and tissues,<sup>26,30–32</sup> the biological consequences remain largely unknown. A number of APA regulators implicated in cancer progression are discussed below, while further studies are required to uncover how APA generates cancer type-specific isoform profiles.

- (1) **CFIm25:** CFIm25 is encoded by the gene '*NUDT21*', and is a subunit of the CFIm complex. CFIm25 binds directly to the UGUA element found in proximity of a subset of PASs. CFIm25 knockdown enhances proximal PAS usage, resulting in the expression of transcripts with shortened 3'UTRs. Interestingly, CFIm25 expression is down-regulated in many different cancer types, which is then accompanied by global 3'UTR shortening.<sup>33,42,127,132,133</sup> In glioblastoma, down-regulation of CFIm25 results in 3'UTR shortening and upregulation of oncogenes such as *CCND1* and *Pak1*,<sup>33,42</sup> enhancing tumorigenic properties and increasing tumor size.<sup>33</sup> Reduced expression of CFIm25 has been observed in hepatocellular carcinoma (HCC) tissue, and forced expression of CFIm25 inhibits HCC cell proliferation, metastasis, and tumorigenesis.<sup>132,134</sup> Mechanistically, CFIm25 knockdown promotes proximal PAS usage in the *CXXC5* and *PSMB2* 3'UTRs, leading to markedly increased expression of both genes and contributing to hepatocellular carcinoma suppression.<sup>134</sup>
- (2) **PCF11:** PCF11, a subunit of the CFIm complex, is also involved in cancer progression. PCF11 controls APA of several hundred transcripts in neuroblastoma, including an operon consisting of multiple neurodifferentiation-regulating genes. PCF11 functions as a key APA regulator of WNT signaling involved in the regulation of proliferation, the cell cycle, apoptosis and neurodifferentiation. Low expression of PCF11 in neuroblastoma has been linked to extensive transcriptome APA and is associated with good prognosis and spontaneous tumor regression, while PCF11 knockdown induces aberrant neurodifferentiation.<sup>135</sup> A recent mechanistic study has shown that PCF11 promotes intronic PAS usage and that down-regulation of PCF11 drives enhanced expression of long genes functioning in cell morphology, adhesion, and migration.<sup>136</sup>
- (3) **hnRNPC:** Fischl et al used subcellular fractionation of colon cancer cells to find that cancer progression is associated with distinct nuclear and cytoplasmic APA profiles. Further studies have shown that hnRNPC overexpression plays a critical role establishing APA profiles characteristic for metastatic colon cancer

cells, while hnRNPC knockdown can reverse the cancer progression-linked APA profile.<sup>137</sup>

### Conclusions and perspective

Our understanding of mRNA APA has benefitted greatly from the recent development of experimental approaches to purify mature RNAs and to study 3' terminal regions through technologies such as 3' end sequencing (3'seq),<sup>138</sup> 3' region extraction and deep sequencing (3'READS),<sup>139,140</sup> Poly(A)-ClickSeq<sup>141</sup> and sequencing APA sites (SAPAS),<sup>142</sup> as well as bioinformatic methods to analyze APA.<sup>27,143–147</sup> Characterization of APA profiles in different cell and tissue types and conditions has highlighted their functional importance in multiple physiological and disease states. Although proteomics methods have identified over 80 proteins that may be involved in mRNA 3'end processing, the roles of most of these factors in APA remain unclear.<sup>11</sup> Furthermore, the impacts of 3'UTR APA on transcript stability and translation of regulated proteins are still controversial. Importantly, while changes in APA are associated with cancer onset and are observed in different cancer cell and tissue models, the mechanisms by which 3'UTR APA is dysregulated in cancer are poorly understood. Understanding the molecular mechanisms of abnormal APA regulation in cancers is critical for identifying new molecular targets for therapeutic intervention. Recent research targeting the APA of the androgen receptor (AR) to coordinately block the expression of resistance-associated pathogenic AR splice variants in prostate cancer is a good example of the cancer therapy applications of these findings.<sup>148</sup> Finally, while manipulating the expression of polyadenylation machinery factors (e.g. CFIm25) can globally regulate 3'UTR length,<sup>33,96</sup> current technologies are not able to directly manipulate the length of a desired 3'UTR. Developing such a novel method may help reveal the causal role of specific 3'UTR APA in regulating gene function and the larger process of tumorigenesis.

### Conflict of Interests

The authors declare no conflict of interest.

### References

1. Crick F. Central dogma of molecular biology. *Nature*. 1970; 227(5258):561–563.
2. Elkon R, Ugalde AP, Agami R. Alternative cleavage and polyadenylation: extent, regulation and function. *Nat Rev Genet*. 2013;14(7):496–506.
3. Gruber AJ, Zavolan M. Alternative cleavage and polyadenylation in health and disease. *Nat Rev Genet*. 2019; 20(10):599–614.
4. Mayr C. What are 3' UTRs doing? *Cold Spring Harb Perspect Biol*. 2019;11(10),a034728.
5. Tian B, Manley JL. Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol*. 2017;18(1):18–30.
6. Tian B, Hu J, Zhang H, Lutz CS. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res*. 2005;33(1):201–212.
7. Beaulieu E, Gautheret D. Identification of alternate polyadenylation sites and analysis of their tissue distribution using EST data. *Genome Res*. 2001;11(9):1520–1526.



8. Venkataraman K, Brown KM, Gilmartin GM. Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes Dev.* 2005;19(11):1315–1327.
9. Takagaki Y, Manley JL. RNA recognition by the human polyadenylation factor CstF. *Mol Cell Biol.* 1997;17(7):3907–3914.
10. Chen F, Wilusz J. Auxiliary downstream elements are required for efficient polyadenylation of mammalian pre-mRNAs. *Nucleic Acids Res.* 1998;26(12):2891–2898.
11. Shi Y, Di Giammartino DC, Taylor D, et al. Molecular architecture of the human pre-mRNA 3' processing complex. *Mol Cell.* 2009;33(3):365–376.
12. Murthy KG, Manley JL. The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes Dev.* 1995;9(21):2672–2683.
13. Kaufmann I, Martin G, Friedlein A, Langen H, Keller W. Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J.* 2004;23(3):616–626.
14. Derti A, Garrett-Engele P, Macisaac KD, et al. A quantitative atlas of polyadenylation in five mammals. *Genome Res.* 2012;22(6):1173–1183.
15. Shi Y. Alternative polyadenylation: new insights from global analyses. *RNA.* 2012;18(12):2105–2117.
16. Wang R, Zheng D, Yehia G, Tian B. A compendium of conserved cleavage and polyadenylation events in mammalian genes. *Genome Res.* 2018;28(10):1427–1441.
17. Reyes A, Huber W. Alternative start and termination sites of transcription drive most transcript isoform differences across human tissues. *Nucleic Acids Res.* 2018;46(2):582–592.
18. Chang JY, Yu WH, Juan HF, Huang HC. Dynamics of alternative polyadenylation in human preimplantation embryos. *Biochem Biophys Res Commun.* 2018;504(4):727–733.
19. Hu W, Li S, Park JY, et al. Dynamic landscape of alternative polyadenylation during retinal development. *Cell Mol Life Sci.* 2017;74(9):1721–1739.
20. Zhou X, Zhang Y, Michal JJ, et al. Alternative polyadenylation coordinates embryonic development, sexual dimorphism and longitudinal growth in *Xenopus tropicalis*. *Cell Mol Life Sci.* 2019;76(11):2185–2198.
21. Sadek J, Omer A, Hall D, Ashour K, Gallouzi IE. Alternative polyadenylation and the stress response. *Wiley Interdiscipl Rev RNA.* 2019;10(5),e1540.
22. Hoque M, Ji Z, Zheng D, et al. Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. *Nat Methods.* 2013;10(2):133–139.
23. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science.* 2008;320(5883):1643–1647.
24. Ji Z, Tian B. Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. *PLoS One.* 2009;4(12),e8419.
25. Mayr C. Evolution and biological roles of alternative 3'UTRs. *Trends Cell Biol.* 2016;26(3):227–237.
26. Mayr C, Bartel DP. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell.* 2009;138(4):673–684.
27. Xia Z, Donehower LA, Cooper TA, et al. Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types. *Nat Commun.* 2014;5, e5274.
28. Hilgers V, Perry MW, Hendrix D, Stark A, Levine M, Haley B. Neural-specific elongation of 3' UTRs during *Drosophila* development. *Proc Natl Acad Sci U S A.* 2011;108(38):15864–15869.
29. Ji Z, Lee JY, Pan Z, Jiang B, Tian B. Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. *Proc Natl Acad Sci U S A.* 2009;106(17):7028–7033.
30. Lee SH, Singh I, Tisdale S, Abdel-Wahab O, Leslie CS, Mayr C. Widespread intronic polyadenylation inactivates tumour suppressor genes in leukaemia. *Nature.* 2018;561(7721):127–131.
31. Singh P, Alley TL, Wright SM, et al. Global changes in processing of mRNA 3' untranslated regions characterize clinically distinct cancer subtypes. *Cancer Res.* 2009;69(24):9422–9430.
32. Fu Y, Sun Y, Li Y, et al. Differential genome-wide profiling of tandem 3' UTRs among human breast cancer and normal cells by high-throughput sequencing. *Genome Res.* 2011;21(5):741–747.
33. Masamha CP, Xia Z, Yang J, et al. CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. *Nature.* 2014;510(7505):412–416.
34. Turner RE, Pattison AD, Beilharz TH. Alternative polyadenylation in the regulation and dysregulation of gene expression. *Semin Cell Dev Biol.* 2018;75:61–69.
35. Li W, You B, Hoque M, et al. Systematic profiling of poly(A)+ transcripts modulated by core 3' end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. *PLoS Genet.* 2015;11(4),e1005166.
36. Tian B, Pan Z, Lee JY. Widespread mRNA polyadenylation events in introns indicate dynamic interplay between polyadenylation and splicing. *Genome Res.* 2007;17(2):156–165.
37. Li QQ, Liu Z, Lu W, Liu M. Interplay between alternative splicing and alternative polyadenylation defines the expression outcome of the plant unique OXIDATIVE TOLERANT-6 gene. *Sci Rep.* 2017;7(1):2052.
38. Berkovits BD, Mayr C. Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. *Nature.* 2015;522(7556):363–367.
39. Hoffman Y, Bublik DR, Ugalde AP, et al. 3'UTR shortening potentiates MicroRNA-based repression of pro-differentiation genes in proliferating human cells. *PLoS Genet.* 2016;12(2), e1005879.
40. Taliadro JM, Vidaki M, Oliveira R, et al. Distal alternative last exons localize mRNAs to neural projections. *Mol Cell.* 2016;61(6):821–833.
41. Lau AG, Irier HA, Gu J, et al. Distinct 3'UTRs differentially regulate activity-dependent translation of brain-derived neurotrophic factor (BDNF). *Proc Natl Acad Sci U S A.* 2010;107(36):15945–15950.
42. Chu Y, Elrod N, Wang C, et al. Nudt21 regulates the alternative polyadenylation of Pak1 and is predictive in the prognosis of glioblastoma patients. *Oncogene.* 2019;38(21):4154–4168.
43. Proudfoot NJ. Ending the message: poly(A) signals then and now. *Genes Dev.* 2011;25(17):1770–1782.
44. Sun Y, Zhang Y, Hamilton K, et al. Molecular basis for the recognition of the human AAUAAA polyadenylation signal. *Proc Natl Acad Sci U S A.* 2018;115(7):E1419–E1428.
45. Clerici M, Faini M, Muckenfuss LM, Aebersold R, Jinek M. Structural basis of AAUAAA polyadenylation signal recognition by the human CPSF complex. *Nat Struct Mol Biol.* 2018;25(2):135–138.
46. Casanal A, Kumar A, Hill CH, et al. Architecture of eukaryotic mRNA 3'-end processing machinery. *Science.* 2017;358(6366):1056–1059.
47. Hu J, Lutz CS, Wilusz J, Tian B. Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. *RNA.* 2005;11(10):1485–1493.
48. MacDonald CC, Redondo JL. Reexamining the polyadenylation signal: were we wrong about AAUAAA? *Mol Cell Endocrinol.* 2002;190(1-2):1–8.

49. Tian B, Graber JH. Signals for pre-mRNA cleavage and polyadenylation. *Wiley Interdiscipl Rev RNA*. 2012;3(3):385–396.
50. Gruber AJ, Schmidt R, Gruber AR, et al. A comprehensive analysis of 3' end sequencing data sets reveals novel polyadenylation signals and the repressive role of heterogeneous ribonucleoprotein C on cleavage and polyadenylation. *Genome Res*. 2016;26(8):1145–1159.
51. Dantone JC, Murthy KG, Manley JL, Tora L. Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA. *Nature*. 1997;389(6649):399–402.
52. Mandel CR, Kaneko S, Zhang H, et al. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature*. 2006;444(7121):953–956.
53. Gil A, Proudfoot NJ. Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit beta-globin mRNA 3' end formation. *Cell*. 1987;49(3):399–406.
54. Cheng Y, Miura RM, Tian B. Prediction of mRNA polyadenylation sites by support vector machine. *Bioinformatics*. 2006;22(19):2320–2325.
55. Murthy KG, Manley JL. Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. *J Biol Chem*. 1992;267(21):14804–14811.
56. Clerici M, Faini M, Aebersold R, Jinek M. Structural insights into the assembly and polyA signal recognition mechanism of the human CPSF complex. *Elife*. 2017;6:e33111.
57. Chan SL, Huppertz I, Yao C, et al. CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3' processing. *Genes Dev*. 2014;28(21):2370–2380.
58. Schonemann L, Kuhn U, Martin G, et al. Reconstitution of CPSF active in polyadenylation: recognition of the polyadenylation signal by WDR33. *Genes Dev*. 2014;28(21):2381–2393.
59. Grozdanov PN, Masoumzadeh E, Latham MP, MacDonald CC. The structural basis of CstF-77 modulation of cleavage and polyadenylation through stimulation of CstF-64 activity. *Nucleic Acids Res*. 2018;46(22):12022–12039.
60. Lackford B, Yao C, Charles GM, et al. Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. *EMBO J*. 2014;33(8):878–889.
61. Shankarling GS, Coates PW, Dass B, Macdonald CC. A family of splice variants of CstF-64 expressed in vertebrate nervous systems. *BMC Mol Biol*. 2009;10:e22.
62. Takagaki Y, Manley JL. Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol Cell Biol*. 2000;20(5):1515–1525.
63. Moreno-Morcillo M, Minvielle-Sebastia L, Fribourg S, Mackereth CD. Locked tether formation by cooperative folding of Rna14p monkeytail and Rna15p hinge domains in the yeast CF IA complex. *Structure*. 2011;19(4):534–545.
64. Bai Y, Auperin TC, Chou CY, Chang GG, Manley JL, Tong L. Crystal structure of murine CstF-77: dimeric association and implications for polyadenylation of mRNA precursors. *Mol Cell*. 2007;25(6):863–875.
65. Yao C, Biesinger J, Wan J, et al. Transcriptome-wide analyses of CstF64-RNA interactions in global regulation of mRNA alternative polyadenylation. *Proc Natl Acad Sci U S A*. 2012;109(46):18773–18778.
66. Brown KM, Gilmartin GM. A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. *Mol Cell*. 2003;12(6):1467–1476.
67. Coseno M, Martin G, Berger C, Gilmartin G, Keller W, Doublet S. Crystal structure of the 25 kDa subunit of human cleavage factor Im. *Nucleic Acids Res*. 2008;36(10):3474–3483.
68. Yang Q, Gilmartin GM, Doublet S. Structural basis of UGUA recognition by the Nudix protein CFIm25 and implications for a regulatory role in mRNA 3' processing. *Proc Natl Acad Sci U S A*. 2010;107(22):10062–10067.
69. Ruegsegger U, Beyer K, Keller W. Purification and characterization of human cleavage factor Im involved in the 3' end processing of messenger RNA precursors. *J Biol Chem*. 1996;271(11):6107–6113.
70. Martin G, Gruber AR, 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*. 2012;1(6):753–763.
71. Zhu Y, Wang X, Forouzmand E, et al. Molecular mechanisms for CFIm-mediated regulation of mRNA alternative polyadenylation. *Mol Cell*. 2018;69(1):62–74.
72. Yang Q, Allard P, Huang M, Zhang W, Clarke HJ. Proteasomal activity is required to initiate and to sustain translational activation of messenger RNA encoding the stem-loop-binding protein during meiotic maturation in mice. *Biol Reprod*. 2010;82(1):123–131.
73. Manley JL, Tacke R. SR proteins and splicing control. *Genes Dev*. 1996;10(13):1569–1579.
74. Ruegsegger U, Blank D, Keller W. Human pre-mRNA cleavage factor Im is related to spliceosomal SR proteins and can be reconstituted in vitro from recombinant subunits. *Mol Cell*. 1998;1(2):243–253.
75. Gruber AR, Martin G, Keller W, Zavolan M. Cleavage factor Im is a key regulator of 3' UTR length. *RNA Biol*. 2012;9(12):1405–1412.
76. Schafer P, Tuting C, Schonemann L, et al. Reconstitution of mammalian cleavage factor II involved in 3' processing of mRNA precursors. *RNA*. 2018;24(12):1721–1737.
77. Kamieniarz-Gdula K, Gdula MR, Panser K, et al. Selective roles of vertebrate PCF11 in premature and full-length transcript termination. *Mol Cell*. 2019;74(1):158–172.
78. Papasaikas P, Valcarcel J. The spliceosome: the ultimate RNA chaperone and sculptor. *Trends Biochem Sci*. 2016;41(1):33–45.
79. Gunderson SI, Polycarpou-Schwarz M, Mattaj JW. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol Cell*. 1998;1(2):255–264.
80. Berg MG, Singh LN, Younis I, et al. U1 snRNP determines mRNA length and regulates isoform expression. *Cell*. 2012;150(1):53–64.
81. Kaida D, Berg MG, Younis I, et al. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature*. 2010;468(7324):664–668.
82. Kuhn U, Gundel M, Knoth A, Kerwitz Y, Rudel S, Wahle E. Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. *J Biol Chem*. 2009;284(34):22803–22814.
83. Tudek A, Schmid M, Jensen TH. Escaping nuclear decay: the significance of mRNA export for gene expression. *Curr Genet*. 2019;65(2):473–476.
84. Jenal M, Elkon R, Loayza-Puch F, et al. The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell*. 2012;149(3):538–553.
85. Sakai Y, Saijo M, Coelho K, Kishino T, Niihara N, Taya Y. cDNA sequence and chromosomal localization of a novel human protein, RBQ-1 (RBBP6), that binds to the retinoblastoma gene product. *Genomics*. 1995;30(1):98–101.
86. Simons A, Melamed-Bessudo C, Wolkowicz R, et al. PACT: cloning and characterization of a cellular p53 binding protein that interacts with Rb. *Oncogene*. 1997;14(2):145–155.
87. Di Giammartino DC, Li W, Ogami K, et al. RBBP6 isoforms regulate the human polyadenylation machinery and modulate expression of mRNAs with AU-rich 3' UTRs. *Genes Dev*. 2014;28(20):2248–2260.
88. Cevher MA, Kleiman FE. Connections between 3'-end processing and DNA damage response. *Wiley Interdiscipl Rev RNA*. 2010;1(1):193–199.

89. Birse CE, Minvielle-Sebastia L, Lee BA, Keller W, Proudfoot NJ. Coupling termination of transcription to messenger RNA maturation in yeast. *Science*. 1998;280(5361):298–301.
90. Pinto PA, Henriques T, Freitas MO, et al. RNA polymerase II kinetics in polo polyadenylation signal selection. *EMBO J*. 2011;30(12):2431–2444.
91. de la Mata M, Alonso CR, Kadener S, et al. A slow RNA polymerase II affects alternative splicing in vivo. *Mol Cell*. 2003;12(2):525–532.
92. Yu L, Rege M, Peterson CL, Volkert MR. RNA polymerase II depletion promotes transcription of alternative mRNA species. *BMC Mol Biol*. 2016;17(1):20.
93. Oktaba K, Zhang W, Lotz TS, et al. ELAV links paused Pol II to alternative polyadenylation in the *Drosophila* nervous system. *Mol Cell*. 2015;57(2):341–348.
94. Chan S, Choi EA, Shi Y. Pre-mRNA 3'-end processing complex assembly and function. *Wiley Interdiscipl Rev RNA*. 2011;2(3):321–335.
95. Nam JW, Rissland OS, Koppstein D, et al. Global analyses of the effect of different cellular contexts on microRNA targeting. *Mol Cell*. 2014;53(6):1031–1043.
96. Brumbaugh J, Di Stefano B, Wang X, et al. Nudt21 controls cell fate by connecting alternative polyadenylation to chromatin signaling. *Cell*. 2018;172(1-2):106–120.
97. Plass M, Rasmussen SH, Krogh A. Highly accessible AU-rich regions in 3' untranslated regions are hotspots for binding of regulatory factors. *PLoS Comput Biol*. 2017;13(4), e1005460.
98. Garneau NL, Wilusz J, Wilusz CJ. The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol*. 2007;8(2):113–126.
99. Gupta I, Clauder-Munster S, Klaus B, et al. Alternative polyadenylation diversifies post-transcriptional regulation by selective RNA-protein interactions. *Mol Syst Biol*. 2014;10(2), e719.
100. Spies N, Burge CB, Bartel DP. 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. *Genome Res*. 2013;23(12):2078–2090.
101. Moqtaderi Z, Geisberg JV, Struhl K. Extensive structural differences of closely related 3' mRNA isoforms: links to Pab1 binding and mRNA stability. *Mol Cell*. 2018;72(5):849–861.
102. Geisberg JV, Moqtaderi Z, Fan X, Ozsolak F, Struhl K. Global analysis of mRNA isoform half-lives reveals stabilizing and destabilizing elements in yeast. *Cell*. 2014;156(4):812–824.
103. Tycowski KT, Shu MD, Steitz JA. Myriad triple-helix-forming structures in the transposable element RNAs of plants and fungi. *Cell Rep*. 2016;15(6):1266–1276.
104. Lee JE, Lee JY, Wilusz J, Tian B, Wilusz CJ. Systematic analysis of cis-elements in unstable mRNAs demonstrates that CUGBP1 is a key regulator of mRNA decay in muscle cells. *PLoS One*. 2010;5(6), e11201.
105. O'Neill KM, Donohue KE, Omelchenko A, Firestein BL. The 3' UTRs of brain-derived neurotrophic factor transcripts differentially regulate the dendritic arbor. *Front Cell Neurosci*. 2018;12, e60.
106. Oe S, Miki H, Nishimura W, Noda Y. Mechanism of the dendritic translation and localization of brain-derived neurotrophic factor. *Cell Struct Funct*. 2016;41(1):23–31.
107. Oliveira MS, Freitas J, Pinto PAB, et al. Cell cycle kinase polo is controlled by a widespread 3' untranslated region regulatory sequence in *Drosophila melanogaster*. *Mol Cell Biol*. 2019;39(15), e00581.
108. Fu Y, Chen L, Chen C, et al. Crosstalk between alternative polyadenylation and miRNAs in the regulation of protein translational efficiency. *Genome Res*. 2018;28(11):1656–1663.
109. An JJ, Gharami K, Liao GY, et al. Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell*. 2008;134(1):175–187.
110. Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM. Localization of ASH1 mRNA particles in living yeast. *Mol Cell*. 1998;2(4):437–445.
111. Niedner A, Edelmann FT, Niessing D. Of social molecules: the interactive assembly of ASH1 mRNA-transport complexes in yeast. *RNA Biol*. 2014;11(8):998–1009.
112. Ephrussi A, Dickinson LK, Lehmann R. Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell*. 1991;66(1):37–50.
113. Neve J, Burger K, Li W, et al. Subcellular RNA profiling links splicing and nuclear DICER1 to alternative cleavage and polyadenylation. *Genome Res*. 2016;26(1):24–35.
114. Chen LL, Carmichael GG. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol Cell*. 2009;35(4):467–478.
115. Loya A, Pnueli L, Yosefzon Y, Wexler Y, Ziv-Ukelson M, Arava Y. The 3'-UTR mediates the cellular localization of an mRNA encoding a short plasma membrane protein. *RNA*. 2008;14(7):1352–1365.
116. Reid DW, Nicchitta CV. Diversity and selectivity in mRNA translation on the endoplasmic reticulum. *Nat Rev Mol Cell Biol*. 2015;16(4):221–231.
117. Taliaferro JM, Lambert NJ, Sudmant PH, et al. RNA sequence context effects measured in vitro predict in vivo protein binding and regulation. *Mol Cell*. 2016;64(2):294–306.
118. Jambhekar A, Derisi JL. Cis-acting determinants of asymmetric, cytoplasmic RNA transport. *RNA*. 2007;13(5):625–642.
119. Yudin D, Hanz S, Yoo S, et al. Localized regulation of axonal RanGTPase controls retrograde injury signaling in peripheral nerve. *Neuron*. 2008;59(2):241–252.
120. Andreassi C, Riccio A. To localize or not to localize: mRNA fate is in 3'UTR ends. *Trends Cell Biol*. 2009;19(9):465–474.
121. Ciolli Mattioli C, Rom A, Franke V, et al. Alternative 3' UTRs direct localization of functionally diverse protein isoforms in neuronal compartments. *Nucleic Acids Res*. 2019;47(5):2560–2573.
122. Ma W, Mayr C. A membraneless organelle associated with the endoplasmic reticulum enables 3'UTR-mediated protein-protein interactions. *Cell*. 2018;175(6):1492–1506.
123. Curinha A, Oliveira Braz S, Pereira-Castro I, Cruz A, Moreira A. Implications of polyadenylation in health and disease. *Nucleus*. 2014;5(6):508–519.
124. Chang JW, Yeh HS, Yong J. Alternative polyadenylation in human diseases. *Endocrinol Metabol*. 2017;32(4):413–421.
125. Lin Y, Li Z, Ozsolak F, et al. An in-depth map of polyadenylation sites in cancer. *Nucleic Acids Res*. 2012;40(17):8460–8471.
126. Lin WH, Gunay C, Marley R, Prinz AA, Baines RA. Activity-dependent alternative splicing increases persistent sodium current and promotes seizure. *J Neurosci*. 2012;32(21):7267–7277.
127. Masamha CP, Wagner EJ. The contribution of alternative polyadenylation to the cancer phenotype. *Carcinogenesis*. 2018;39(1):2–10.
128. Xue Z, Warren RL, Gibb EA, et al. Recurrent tumor-specific regulation of alternative polyadenylation of cancer-related genes. *BMC Genomics*. 2018;19(1), e536.
129. Wang L, Hu X, Wang P, Shao ZM. The 3'UTR signature defines a highly metastatic subgroup of triple-negative breast cancer. *Oncotarget*. 2016;7(37):59834–59844.
130. Wang L, Hu X, Wang P, Shao ZM. Integrative 3' untranslated region-based model to identify patients with low risk of axillary lymph node metastasis in operable triple-negative breast cancer. *The Oncologist*. 2019;24(1):22–30.

131. Park HJ, Ji P, Kim S, et al. 3' UTR shortening represses tumor-suppressor genes in trans by disrupting ceRNA crosstalk. *Nat Genet.* 2018;50(6):783–789.
132. Sun M, Ding J, Li D, Yang G, Cheng Z, Zhu Q. NUDT21 regulates 3'-UTR length and microRNA-mediated gene silencing in hepatocellular carcinoma. *Cancer Lett.* 2017;410:158–168.
133. Zhu ZJ, Huang P, Chong YX, et al. MicroRNA-181a promotes proliferation and inhibits apoptosis by suppressing CFIm25 in osteosarcoma. *Mol Med Rep.* 2016;14(5):4271–4278.
134. Tan S, Li H, Zhang W, et al. NUDT21 negatively regulates PSMB2 and CXXC5 by alternative polyadenylation and contributes to hepatocellular carcinoma suppression. *Oncogene.* 2018;37(35):4887–4900.
135. Ogorodnikov A, Levin M, Tattikota S, et al. Transcriptome 3' end organization by PCF11 links alternative polyadenylation to formation and neuronal differentiation of neuroblastoma. *Nat Commun.* 2018;9(1),e5331.
136. Wang R, Zheng D, Wei L, Ding Q, Tian B. Regulation of intronic polyadenylation by PCF11 impacts mRNA expression of long genes. *Cell Rep.* 2019;26(10):2766–2778.
137. Fischl H, Neve J, Wang Z, et al. hnRNPC regulates cancer-specific alternative cleavage and polyadenylation profiles. *Nucleic Acids Res.* 2019;47(14):7580–7591.
138. Lianoglou S, Garg V, Yang JL, Leslie CS, Mayr C. Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes Dev.* 2013;27(21):2380–2396.
139. Zheng D, Liu X, Tian B. 3'READS+, a sensitive and accurate method for 3' end sequencing of polyadenylated RNA. *RNA.* 2016;22(10):1631–1639.
140. Hoque M, Li W, Tian B. Accurate mapping of cleavage and polyadenylation sites by 3' region extraction and deep sequencing. *Methods Mol Biol.* 2014;1125:119–129.
141. Routh A, Ji P, Jaworski E, Xia Z, Li W, Wagner EJ. Poly(A)-ClickSeq: click-chemistry for next-generation 3-end sequencing without RNA enrichment or fragmentation. *Nucleic Acids Res.* 2017;45(12),e112.
142. Wang X, Li M, Yin Y, et al. Profiling of alternative polyadenylation sites in luminal B breast cancer using the SAPAS method. *Int J Mol Med.* 2015;35(1):39–50.
143. Ha KCH, Blencowe BJ, Morris Q. QAPA: a new method for the systematic analysis of alternative polyadenylation from RNA-seq data. *Genome Biol.* 2018;19(1),e45.
144. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol.* 2016;34(5):525–527.
145. Ye C, Long Y, Ji G, Li QQ, Wu X. APATrap: identification and quantification of alternative polyadenylation sites from RNA-seq data. *Bioinformatics.* 2018;34(11):1841–1849.
146. Chang JW, Zhang W, Yeh HS, et al. An integrative model for alternative polyadenylation, IntMAP, delineates mTOR-modulated endoplasmic reticulum stress response. *Nucleic Acids Res.* 2018;46(12):5996–6008.
147. Arefeen A, Liu J, Xiao X, Jiang T. TAPAS: tool for alternative polyadenylation site analysis. *Bioinformatics.* 2018;34(15):2521–2529.
148. Van Etten JL, Nyquist M, Li Y, et al. Targeting a single alternative polyadenylation site coordinately blocks expression of androgen receptor mRNA splice variants in prostate cancer. *Cancer Res.* 2017;77(19):5228–5235.