

## Cleavage and polyadenylation: Ending the message expands gene regulation

Jonathan Neve\*, Radhika Patel , Zhiqiao Wang , Alastair Louey , and André Martin Furger 

Department of Biochemistry, University of Oxford, Oxford, United Kingdom

### ABSTRACT

Cleavage and polyadenylation (pA) is a fundamental step that is required for the maturation of primary protein encoding transcripts into functional mRNAs that can be exported from the nucleus and translated in the cytoplasm. 3' end processing is dependent on the assembly of a multiprotein processing complex on the pA signals that reside in the pre-mRNAs. Most eukaryotic genes have multiple pA signals, resulting in alternative cleavage and polyadenylation (APA), a widespread phenomenon that is important to establish cell state and cell type specific transcriptomes. Here, we review how pA sites are recognized and comprehensively summarize how APA is regulated and creates mRNA isoform profiles that are characteristic for cell types, tissues, cellular states and disease.

### ARTICLE HISTORY

Received 9 January 2017  
Revised 2 March 2017  
Accepted 9 March 2017

### KEYWORDS

3' end processing; alternative cleavage and polyadenylation; gene expression

### Introduction

To express genes, the information present in the DNA blueprint is transcribed by RNA polymerases into RNA molecules. These initial or primary transcripts are often subjected to modifications and processing reactions that mature them into functional molecules. While the processing of some noncoding RNAs such as rRNA and tRNAs are also essential in prokaryotes, extensive processing of protein encoding pre-mRNA transcripts is a unique feature of eukaryotic gene expression.<sup>1</sup>

In eukaryotes, 3 major processing reactions are responsible for the maturation of pre-mRNA molecules into mRNAs that are competent for nuclear cytoplasmic export and subsequent translation.<sup>2</sup> Firstly, the 5' ends of RNAs are modified by the addition of a cap structure in the form of a guanosine molecule that is enzymatically attached to the first nucleotide of the transcribed mRNA by an unusual 5'-5' triphosphate linkage and subsequently modified by methylation. The genes that encode proteins are mostly not continuous, but separated into alternating exons (coding regions) and introns, which are noncoding regions that need to be removed. This occurs during splicing, where a large multi-protein-RNA complex, known as the spliceosome, directs precise excision of introns from the primary transcript and fuses the exons together. This generates an mRNA with a continuous open reading frame (ORF) that serves as a template for protein synthesis in the cytoplasm. In order for the transcript to be exported to the cytoplasm, the transcript needs to be detached from RNA Polymerase, which occurs by a 2-step cleavage and polyadenylation reaction. This forms the typical uniform polyadenylated 3' end of almost all of the protein encoding nuclear mRNAs.

While splicing has long been recognized to play an important role in the regulation of gene expression in eukaryotes, the

scope for gene regulation at the level of cleavage and polyadenylation has only recently been fully recognized.

### Core cleavage and polyadenylation sequence elements

With the exception of replication dependent histone genes, metazoan protein encoding mRNAs contain a uniform 3' end consisting of a stretch of adenosines. This poly(A) tail is synthesized in a template independent fashion and is the result of an RNA processing reaction, where a multiprotein complex assembles on specific sequences on the pre-mRNAs, called the cleavage and polyadenylation signals (pA signals). pA signals consist of sequences that flank either side of where the pre-mRNA is endonucleolytically cleaved and subsequently polyadenylated (Fig. 1). The core pA signal is a bipartite sequence element that constitutes an A-rich hexameric sequence, found 21 nucleotides upstream of the cleavage site<sup>3</sup> and a U and/or G/U rich sequence, located 10–30 nucleotides downstream of the cleavage site. Whereas the upstream hexamer motif for most (70–75%) pA signals conforms to the canonical A[A/U]UAAA hexamer,<sup>4,5</sup> the nucleotide composition of the downstream sequence element (DSE) can vary significantly between pA signals. Although no clear consensus can be defined,<sup>6</sup> often a GU rich sequence followed by a stretch of 3 or more uridines is present within a 40 nucleotide wide window downstream of the cleavage site.<sup>7</sup>

### Core sequences upstream of the cleavage site

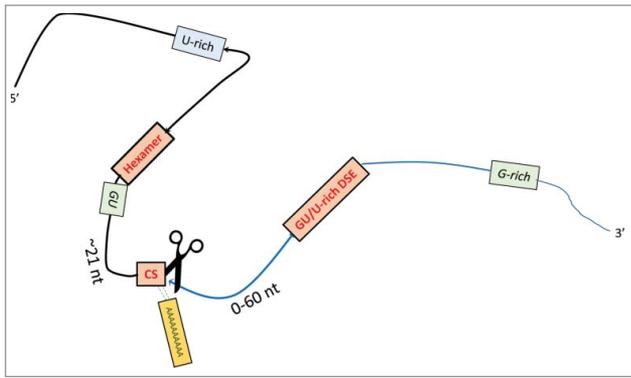
The importance of the nucleotide composition in the hexamer sequence element is underpinned by numerous *in vivo* and *in vitro* experimental examples. Single nucleotide changes in the

**CONTACT** André Martin Furger  [andre.furger@bioch.ox.ac.uk](mailto:andre.furger@bioch.ox.ac.uk)  Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, United Kingdom.

 Supplemental data for this article can be accessed in the publisher's website.

\*These authors contributed equally to this work.

Published with license by Taylor & Francis Group, LLC © Jonathan Neve, Radhika Patel, Zhiqiao Wang, Alastair Louey, and André Martin Furger  
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.



**Figure 1.** The cis-elements that define pA sites. The cleavage and polyadenylation machinery relies on key cis elements to mediate 3' end processing. Canonical cis elements include the A[A/U]UAAA hexamer and its variants which lie ~21 nucleotides upstream of the cleavage site (CS) and a downstream less well defined GU/U-rich element. Additional auxiliary elements may be positioned upstream and/or downstream of the cleavage site and are often U, GU and or G-rich.

A[A/U]UAAA sequence, or its entire removal,<sup>8-10</sup> have proven to severely impair cleavage efficiency. These laboratory based observations are further strengthened by the naturally occurring polymorphisms such as found in the human globin genes,<sup>11-14</sup> which are associated with the thalassemia phenotype. However, these single nucleotide polymorphisms (SNPs) rarely occur in the pA hexamer,<sup>15</sup> further emphasizing the importance of the hexamer nucleotide composition. However, naturally occurring **non-canonical variants**, with a deviation of one or more nucleotides from the A[A/U]UAAA hexamer sequence consensus, represent about 10-20% of all hexamer signals.<sup>4,5,12-14,16</sup>

### Core sequences downstream of the cleavage site

In contrast to the hexamer, the DSE has long been considered to be much more tolerant to single nucleotide changes.<sup>17</sup> However, at least for a subset of genes, in particular those with degenerated hexamer sequences, the opposite is true: while mutations in the hexamer had little effect, base changes in the DSE reduced 3' end processing significantly.<sup>18</sup> That the DSE can be a decisive element in regulating efficiency of cleavage is further supported by clinically relevant examples where changes in the DSE force a gain of function and increase the processing capacity of a particular pA site.<sup>12,19</sup> In addition, knockdown of CstF64, the component of the 3' end processing machinery known to interact with the DSE (see below), affects usage of pA sites that are characterized by a non-canonical hexamer and a GUKKU "type" DSE.<sup>20</sup> Furthermore, for a subset of genes that have pA sites with degenerated DSEs, CstF64 is not required for 3' end processing. The mRNAs of these genes are polyadenylated by Star-PAP, a non-canonical poly(A) polymerase.<sup>21</sup> These observations suggest that for some pA sites, the DSE may represent the critical core element that drives cleavage efficiency and directs how the 3' end machinery is assembled at the pA site.

### The cleavage site

The point of cleavage occurs between the hexamer and DSE. In mammals, CA and UA appear to be the most frequent dinucleotides that precede the actual site of cleavage,<sup>22</sup> however, the

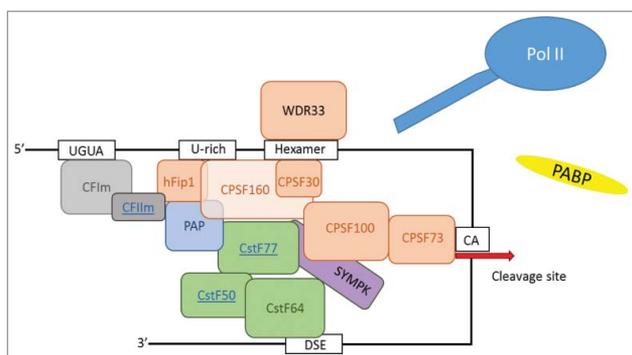
actual site where cleavage occurs is known to be heterogeneous. Nucleotide composition at the cleavage site has a complex impact on 3' end processing. Initial *in vitro* based studies revealed that mutations of the dinucleotides at the cleavage site generally resulted in relatively moderate effects on cleavage efficiency. Conversely, a prominent SNP occurring at a 1% frequency in the Caucasian population, altering the cleavage site of the prothrombin gene from CG to CA, significantly increases cleavage efficiency of the pre-mRNA, and has been linked to a higher tendency toward thrombophilia.<sup>23</sup>

### Auxiliary elements

In addition to the core sequence elements, many pA sites possess additional auxiliary sequences that influence their overall strength. These are diverse in location and nucleotide composition, but in most cases feature U/G rich sequences.<sup>7</sup> The presence of these auxiliary elements correlates with deviation from the consensus A[A/U]UAAA hexamer, perhaps compensating for the loss of core sequence integrity.<sup>18</sup> These "processing enhancers" can be found both upstream and downstream of the cleavage sites.

G-rich auxiliary sequences are some of the best characterized downstream enhancers that may fold into G-quadruplex structures. Several viral and mammalian pA sites have been described where these G-rich sequences significantly influence cleavage efficiency at the pA site<sup>6</sup> by interacting with factors, such as hnRNP, that stimulate processing.<sup>24,25</sup> Because the G-rich sequences can exert their influence over distances that span several hundred bases,<sup>24,26,27</sup> the true extent to which pA sites depend on these distant enhancers may be undervalued.

U-rich auxiliary regions are frequently found upstream and downstream of the cleavage site.<sup>3,7</sup> The upstream U-rich sequences can be divided into 2 categories. The first category represents U-rich elements that are located between the hexamer and the actual cleavage site.<sup>7,28</sup> They can function as binding sites for the CPSF subunit Fip1, which can stimulate the polyadenylation step<sup>29</sup> and control pA site selection.<sup>30,31</sup> The second category includes U-rich sequences that are positioned in the UTR upstream of the hexamer. These upstream elements (USEs) were first identified in several viral transcripts including SV40,<sup>32</sup> Cauliflower Mosaic Virus (CaMV)<sup>33</sup> and HIV-1.<sup>34</sup> In the HIV-1 transcript the USE, identified between 56 and 93 nucleotides upstream of the hexamer, was found to interact with CPSF160 and aid its interaction with the poly(A) signal, proving to be critical for the cleavage reaction.<sup>35</sup> USEs generally have no consensus sequences, although tend to be U-rich.<sup>17</sup> A number of USEs have now been identified in transcripts of the human genome. In the lamin B2 and C2 complement genes, auxiliary USEs are required for interaction with PTB and CstF64 respectively, which is critical for the cleavage and polyadenylation reaction.<sup>36,37</sup> A USE controls the usage of the poly(A) sites in the prothrombin gene in a stress response,<sup>38</sup> and efficiency of polyadenylation in collagen genes has also been shown to be modulated by USEs.<sup>39</sup> In addition, the usage of the proximal poly(A) site of the cyclo-oxygenase 2 gene is controlled by 3 U-rich USEs that bind a protein complex to recruit the core polyadenylation factors.<sup>40,41</sup>



**Figure 2.** The core factors of the cleavage and polyadenylation complex. There are more than 80 proteins associated with the cleavage and polyadenylation machinery but fewer than 20 factors are considered to build the core of the processing complex.<sup>47</sup> The major components are made up of multi-subunit factors including the cleavage and polyadenylation specificity factor CPSF (WDR33, hFip1, CPSF160, CPSF100, CPSF70, CPSF30); the cleavage stimulatory factor CstF (CstF77, CstF64, CstF50), the CFI (CFIm65, CFIm25) and CFII (~15 subunits). The core factors involved in cleavage and polyadenylation, and the cis elements to which they bind are outlined here. Details of the individual factors are given in the text.

A UGUA motif has been identified as a USE, which is bound by the cleavage factor I (CFI) complex. UGUA mediated CFIm recruitment to the pre-mRNA can direct cleavage and polyadenylation at non-canonical sites, by anchoring Fip1 and the poly(A) polymerase (PAP) to the pre-mRNA.<sup>42,43</sup> More recent whole transcriptome based analyses identified a more global role for this interaction and recognized CFI as a key regulator of pA site choice during alternative cleavage and polyadenylation (APA) for a significant number of genes.<sup>44,45</sup> Finally, some of the U-rich elements have also been shown to mediate an interaction with hnRNPC similarly influencing pA site selection during APA.<sup>3</sup>

**AUA auxiliary element.** A subset of mRNAs are polyadenylated by the non-canonical poly(A) polymerase Star-PAP (see below), which binds an AUA motif in the 3'UTR of its target pre-mRNAs.<sup>21</sup>

### The core poly(A) machinery

In mammals there are more than 80 proteins associated with the 3' end processing machinery,<sup>46</sup> but this can be condensed to around 20 proteins that constitute the core factors<sup>47</sup> (Fig. 2). Part of this core machinery includes 4 multi-subunit protein complexes which are highly conserved: the cleavage and polyadenylation specificity factor (CPSF), the cleavage stimulation factor (CstF), and mammalian cleavage factors I and II (CFIm and CFII). Additional core factors include symplekin, the poly(A) polymerase (PAP), nuclear poly(A) binding protein (PABPN1) and the C-terminal domain (CTD) of the largest subunit of RNA polymerase II (Pol II). The process of 3' end formation is initiated by the recognition of the respective cis-elements on the nascent transcripts by CPSF and CstF complexes in a cooperative manner.<sup>48</sup> Depending on the pA site, this initial step is supported and often dependent on the interaction of other core factors (for example CFI)<sup>43</sup> and also auxiliary factors (for example hnRNP H).<sup>24</sup> After a functional complex is assembled cleavage and subsequent polyadenylation of the pre-mRNA can commence.

### CPSF

The CPSF protein complex features 6 protein subunits; WDR33, CPSF30, CPSF160, hFip1, CPSF100 and CPSF73, which are required for cleavage of the pre-mRNA. In contrast, only 4 CPSF subunits, CPSF160, CPSF30, WDR33 and hFip1 together with PAP are sufficient to direct AAUAAA dependent poly(A) tail addition *in vitro*.<sup>49</sup> The different dependencies on the CPSF subunits raises the intriguing possibility that different sub-complexes may be specific for the cleavage or the polyadenylation steps.<sup>49</sup> This idea is supported by the finding that different CPSF sub-complexes have been identified that cater either for pA or histone 3' end processing.<sup>50</sup>

**WDR33 (WDR33)** is a large 146 kDa protein, and despite the essential role its homolog Pfs2p plays in 3' end formation in yeast,<sup>51</sup> its role during mammalian cleavage and polyadenylation has only recently been uncovered. The mammalian WDR33 was first shown to be essential in cleavage and polyadenylation *in vitro*,<sup>52</sup> and has subsequently been identified as the CPSF subunit that directly binds the AAUAAA hexamer.<sup>49,53</sup> Its importance for pA site recognition is further underpinned by PAR-CLIP data, which shows a strong interaction between WDR33 and the pre-mRNA directly upstream (−16 to −18 nt) of cleavage sites,<sup>49</sup> close to the region where hexamers peak in transcripts.<sup>3</sup> It is currently unclear how its specific interaction with AAUAAA is achieved, but the WD40 domain, which is principally associated with protein-protein interactions, also has RNA-binding properties.<sup>54</sup>

**CPSF30 (CPSF4)** is the smallest subunit of the CPSF complex. The involvement of a 30 kDa protein in the recognition of AAUAAA has long been recognized,<sup>55</sup> but it took more than 25 years to show that this protein, together with WDR33, is responsible for the specific interaction of CPSF with the AAUAAA hexamer in mammals.<sup>53</sup> Interestingly, CPSF30 has also been shown to play a central role in pA site selection in plants.<sup>56,57</sup> The RNA-binding properties of CPSF30 reside in the 5 zinc finger (ZF) domains, primarily mediated by ZF2 and ZF3.<sup>53,58</sup> The binding of CPSF30 to RNA is required for the entire CPSF complex to interact with the RNA substrate.<sup>53,59</sup> Notably, the CPSF30-AAUAAA interaction is specifically inhibited by the influenza protein NS1A, that blocks ZF2 and ZF3, and so globally abrogates host 3' end processing.<sup>60</sup> These observations clearly highlight the central role of CPSF30 in pA site recognition.

**CPSF160 (CPSF1)** was considered to be the major component that mediates the binding of the CPSF complex to AAUAAA. This was primarily based on the observation that a 160 kDa protein can be specifically cross linked to AAUAAA-containing RNA substrates<sup>61</sup> and the Yhh1/Cft1p, the yeast homolog of CPSF160, also bound RNA.<sup>62</sup> Furthermore, pull-down assays confirmed that recombinant CPSF160 interacts with AAUAAA-containing RNAs. However, as mentioned above, it is now clear that AAUAAA-specific binding of the CPSF complex falls to WDR33 and CPSF30 and thus the precise role of CPSF160 for the interaction with the pre-mRNA needs to be re-evaluated. Nevertheless, CPSF160 is a critical component of the cleavage and polyadenylation machinery and establishes important protein-protein interactions. These include interactions

between CPSF and CstF by directly contacting the CstF77 subunit. This interaction may be integral to the cooperative nature of pA site recognition. In addition, CPSF160 also interacts functionally with the PAP.<sup>63</sup>

**hFip1 (FIP1L1)**, the human factor interacting with PAP, was discovered based on the sequence similarity to the yeast polyadenylation factor Fip1p. True to its name, Fip1, and hFip1, interact with and recruit PAP to the pA site where it can simultaneously bind to U-rich regions that are often located either upstream or downstream of the pA hexamer. The RNA-binding activity of hFip1 lies within the arginine-rich C-terminus.<sup>29</sup> A role in pA site recognition is further supported by the fact that hFip1 levels change in stem cells compared with differentiated cells. This directs APA during stem cell self-renewal. Higher levels of hFip1 in ESCs and iPSCs promote the recognition of weaker, hFip1 dependent pA sites.<sup>30</sup>

**CPSF100 (CPSF2) and CPSF73 (CPSF3)**, which both contain metallo- $\beta$ -lactamase and  $\beta$ -CASP domains, form a heterodimer in mammals.<sup>64</sup> The precise role of CPSF100 for cleavage and polyadenylation in mammals is still unclear. There is however evidence that CPSF100 is involved in THOC5 (a member of the TREX export complex) mediated control of 3' end processing of immediate early genes. THOC5 recruits CPSF100 to the 3' end of immediate early genes that are regulated by THOC5.<sup>65</sup> The CASP domain in CPSF73 provides the zinc-dependent endonuclease activity during 3' end formation for both the cleavage and polyadenylation<sup>66</sup> and the histone 3' end processing machineries.<sup>67</sup> While they play a central role during cleavage, both proteins are superfluous for the polyadenylation step.

## CstF

**CstF** is the component of the cleavage and polyadenylation machinery that specifically recognizes the DSE region in pA sites. It is a heterotrimeric protein complex containing the subunits CstF77, CstF64 and CstF50.<sup>48,68</sup> The CstF complex is essential for the cleavage reaction but is not required for the polyadenylation step.<sup>69,70</sup>

**CstF64 (CSTF2)** mediates the binding to the DSE in the pre-mRNA.<sup>71</sup> The interaction interface resides in an RNA recognition motif (RRM) at its N-terminus, which specifically recognizes U/GU-rich regions downstream of the cleavage sites. How CstF64 is able to recognize diverse DSEs in pre-mRNAs remains uncertain, but consecutive Us are required for strong CstF-GU interactions. Interestingly, RNA binding also induces structural changes in CstF64 that may be critical to facilitate assembly of larger complexes.<sup>72</sup> A global analysis that mapped transcriptome wide CstF64-RNA interactions *in vivo* suggests a role for CstF64 in modulating the selection of pA sites during APA.<sup>73</sup> This promotes the usage of non-canonical pA sites that feature a GUKKU motif in their DSE.<sup>20</sup>

**CstF77 (CSTF3)** bridges the CstF64 and CstF50 subunits<sup>74</sup> and thus plays a key role for the assembly of the CstF complex. CstF77 appears to be an elongated dimer, suggesting that CstF may form a hexameric complex in cells.<sup>75</sup> CstF77 also appears to be critical for the establishment of protein interactions between CstF and other pA complexes. Cross linking experiments demonstrate a strong interaction between CPSF160 and

CstF77, which may provide the basis for the observed CPSF-CstF cooperative RNA binding during the assembly process.<sup>63</sup>

**CstF50 (CSTF1)**, like CstF77, exists as a homodimer<sup>76</sup> providing further evidence for the hexameric nature of functional CstF complexes. CstF50 contains 7 repeats of the WD40 motif, which mediates protein-protein interactions with CstF77 and BRCA1 associated RING domain 1 (BARD1).<sup>77</sup> The latter interaction is inhibitory and believed to suppress erroneous 3' end formation possibly during transcription coupled DNA repair.<sup>78</sup> CstF50 also interacts with the CTD of Pol II and this interaction may be established via the N-terminal part of CstF50.<sup>79</sup>

**CFIm** is a complex is made up of 4 polypeptides: **CFIm25 (CPSF5/NUDT21)**, **CFIm68 (CPSF6)**, **CFIm59 (CPSF7)** and **CFIm72 (alternative spliced isoform of CFIm68)**. It has been suggested that different forms of CFIm complex exist *in vivo* including CFIm25-68, CFIm25-59 and possibly CFIm25-72. As shown by SELEX experiments<sup>42</sup> and site directed mutagenesis,<sup>43</sup> CFIm recognizes a UGUAN motif in RNA through the RRM in CFIm25. The RRM in CFIm68 is critical for the interaction with CFIm25 and enhances RNA-binding of the complex. CFIm68 also facilitates RNA looping, which may play a role in the regulation of APA.<sup>80,81</sup> CFIm has originally been shown to stabilize the binding of CPSF complex to the pre-mRNA<sup>82</sup> and then emerged as a critical factor for the recognition of pA sites that lack the A[A/U]UAAA hexamer.<sup>43</sup> To date, CFI is arguably the best characterized factor that regulates APA at the point of cleavage. This is demonstrated by numerous examples where a depletion of CFIm results in the preferred usage of the proximal pA sites.<sup>20,44,45,83</sup> CFIm has also been implicated in regulating alternative usage of non-canonical pA sites during spermatogenesis.<sup>84</sup> There are also contacts between CFIm68 and THOC5, a component of the transcription export complex (TREX). THOC5 appears to be important for the recruitment of CFIm68 to distal pA sites.<sup>85</sup>

**CFII<sub>m</sub>**, mammalian CFII, is only required for the cleavage step and is the least well characterized complex of the core component. To date, this complex has only been partly purified and can be separated into 2 fractions: CFII<sub>Am</sub> and CFII<sub>Bm</sub>. The CFII<sub>Am</sub> fraction contains the essential components for the cleavage reaction while CFII<sub>Bm</sub> is not required but has a stimulatory role for cleavage. One of the 15 polypeptides associated with CFII<sub>Am</sub>, **hClp1** has been shown to interact with CFIm and CPSF by immunoprecipitation experiments, suggesting that it bridges these 2 during the cleavage process.<sup>86</sup> An additional component of CFII is protein 1 of cleavage factor 1, **hPcf11**, which has been shown to be critical in the control of Pol II transcription termination, by mediating the degradation of the polymerase associated 3' RNA product after cleavage at the pA site.<sup>87</sup> More recently, hPcf11 has also been implicated in regulating APA.<sup>31</sup>

**PAP (PAPOLA)** is a monomer and responsible for the addition of the 3' polyadenosine tail to a newly synthesized pre-mRNA molecule<sup>88</sup> by catalyzing the chemical reaction  $\text{ATP} + \text{RNA-3'OH} \rightarrow \text{pyrophosphate} + \text{RNAPa-3'OH}$ .<sup>89,90</sup> PAP adds 200–250 adenosines in a template-independent manner to the 3' end of the cleaved pre-mRNA. PAP is recruited to the 3' end processing complex by CPSF through interactions mediated by CPSF160 and hFip1.<sup>29</sup> There are several PAP

isoforms present in human cells which can chiefly be separated into canonical PAPs and non-canonical PAPs. Most RNAs are polyadenylated by canonical PAPs. Non-canonical PAPs, among other functions, are associated with mitochondrial mRNA polyadenylation (hmtPAP), cytoplasmic polyadenylation (hGld2) and miRNA biogenesis. A prominent non-canonical PAP called Star-PAP, controls the polyadenylation of a subset of mRNAs that encode for proteins that are associated with DNA damage induced apoptosis and stress responses.<sup>91</sup>

**PABPN1 (PABPN1)** has been shown to contain a single RRM and an arginine rich CTD.<sup>92</sup> The rate at which PAP adds adenosine nucleotides is dependent on the presence of PABPN1. The first few nucleotides added by PAP are added with low efficiency until the short polyadenosine tail is bound by PABPN1, which accelerates the rate of adenosine addition by PAP.<sup>93</sup> While PAP adds 200–250 adenosine nucleotides to the 3′ end of the mRNAs,<sup>94</sup> the median length of the poly(A) tail in cellular mRNAs is much shorter, between 50 to 100 nucleotides long due to the action of cytoplasmic deadenylases.<sup>95</sup> In addition to its primary function, PABPN1 has also been implicated in controlling APA.<sup>96,97</sup>

**Symplekin (SYMPK)** is a protein that forms a high-molecular weight complex with CPSF and CstF by interacting with CPSF73, CPSF100 and CstF64. It is suggested to serve as a scaffold for recruiting other factors to the cleavage and polyadenylation complex.<sup>98</sup> It also participates in the assembly of a processing complex that matures histone mRNA 3′ ends, which do not undergo polyadenylation. Symplekin has also been found to form a complex with heat shock transcription factor 1 (HSF1) after stress treatment. This complex may be critical for the recruitment of the 3′ end processing machinery to the heat shock protein *HSB* pre-mRNA during stress and so safeguard its expression.<sup>99</sup> Symplekin appears to provide a scaffold around which different CPSF sub-complexes are assembled. Perhaps these different sub-complexes provide the necessary scope and flexibility to the 3′ end machinery that is required to associate with pA sequences that differ widely in their architecture.

**Pol II CTD (RBP1)** is the key platform that couples pre-mRNA processing to transcription and it is required for efficient cleavage at the pA site.<sup>100</sup> The human Pol II CTD features 21 consensus YSPTSPS and 31 non-consensus heptad repeats<sup>101</sup> which form and present a dynamic interaction platform for a variety of pre-mRNA processing components including cleavage and polyadenylation factors such as Pcf11, CstF77 and CstF50.

### The assembly of a functional cleavage and polyadenylation complex

The initiating step of the 3′ end processing complex assembly (Fig. 2) is the coordinated recognition of the hexamer signal AAUAAA and the DSE on the nascent RNA by CPSF and CstF. The recruitment of CPSF is central, as it constitutes the core processing complex required for both the cleavage and the subsequent polyadenylation reaction. It binds the RNA directly by the association of WDR33 and CPSF30 with the hexamer. Whether the 2 subunits bind to pre-mRNAs concurrently or in a time and/or pA site type dependent manner remains

unknown. It may well be that different CPSF complexes provide a flexible machinery that can recognize different types of pA signals. Unlike CPSF, the CstF complex is essential only for the cleavage reaction. CstF associates with di-uridine pockets in the U/GU rich DSE via its RRM in CstF64.<sup>71</sup> CstF77 is a critical component that bridges the CPSF and CstF components, facilitating the cooperative binding.<sup>63</sup> With CPSF and CstF anchored to the pA site, additional factors join and complete the assembly of a functional 3′ end processing complex. The complex can then catalyze a 2-step reaction comprising of an endonucleolytic cleavage at the cleavage site and the polyadenylation of the 3′ end of the 5′ cleavage product. RNA cleavage occurs between the hexamer and the DSE, generally within a window of 20 nucleotides, either side of these core sequences.

The speed of assembly of a functional cleavage and polyadenylation complex is dependent on the strength of the pA site, but is generally achieved within about 10 seconds.<sup>102</sup> The absolute strength of a given pA site is highly complex and depends on the sequence architecture of the pA site (see above) and the relative availability of core and auxiliary promoting or inhibitory factors.

### Interconnection between pA cleavage, capping and splicing

The assembly process of the functional machinery is further complicated by the interconnection between 3′ end processing with capping, pre-mRNA splicing and transcription. The relationship between capping and 3′ end formation was first suggested after capped RNAs proved to be better substrates for 3′ end processing in nuclear extracts.<sup>103</sup> These initial observations were confirmed and shown to be mediated by the cap binding complex, CBC.<sup>104</sup>

The relationship between splicing and 3′ end formation is extensive and principally relies on direct interactions between splicing and cleavage and polyadenylation components. The positive influence of splicing factors on 3′ end formation was first reported over 25 years ago and has since been extensively documented. Intron insertion into a replication dependent histone gene reporter is one example of this interconnection, where a splicing event in the histone pre-mRNA is concomitant with the activation of cryptic pA sites. As a result, cleavage and polyadenylation in the spliced transcripts is favored over U7 snRNP dependent 3′ end formation,<sup>105</sup> which is typical for replication dependent histone genes (reviewed in Romeo and Schümperli, 2016).<sup>106</sup> Splicing has not only been shown to activate cryptic pA sites but generally stimulates cleavage and polyadenylation of the downstream pA sites.<sup>107–110</sup> This stimulation is achieved by interactions between components of the splicing machinery that associate with the 3′ splice sites (3SS) and several poly(A) factors notably between the 65 kDa subunit of U2AF, PAP and CFIm59.<sup>111–113</sup> In addition, physical interactions of CPSF components with several subunits of the U2 snRNP, including SF3b155, SF3b130 and SF3b49 have been found.<sup>114</sup> Correspondingly, mutations of U2 snRNA-binding sites significantly reduced the cleavage efficiency in reporter genes. The stimulatory effect through these interactions is mutual as mutations of the hexamer, can inhibit terminal intron removal<sup>109</sup> and the depletion of CPSF100 impaired

upstream splicing in an *in vitro* coupled splicing and 3' processing system.<sup>114</sup> CPSF and symplekin have been implicated in the promotion of alternative splicing on a global scale<sup>115</sup> which further emphasizes the close connections of splicing and 3' end formation.

While splicing components associated with the 3SS appear to stimulate 3' end formation, the opposite is generally true for splicing components that bind to the 5SS, notably the U1 snRNP. The suppression of pA sites by U1 snRNP has proven to be instrumental for the inhibition of late gene expression in the bovine papilloma type 1 virus<sup>116-118</sup> and in HIV-1, where the U1 snRNP bound to 5SS acts as a suppressor of upstream pA site usage.<sup>119,120</sup> In the latter example, suppression of the upstream pA site is crucial to enable transcription of viral protein-encoding mRNA and the genomic RNA. This U1 snRNP-mediated type of suppression of 3' end formation is not restricted to viruses but was suggested to prevent premature transcription termination in polycistronic transcription units at a global scale in *Caenorhabditis elegans* by suppressing the recognition of the pA sites in upstream positioned genes in polycistronic mRNAs by Pol II associated poly(A) factors.<sup>121</sup> Importantly, the suppressive role of U1 snRNP for 3' end formation emerged as a general mechanism by which cryptic pA sites in eukaryotic introns are suppressed globally, thereby protecting the cell from aberrant pA usage and premature transcription termination.<sup>122</sup> Physiologically, U1 snRNP levels are very high which may ensure that levels never fall below a critical threshold that would result in the activation of these cryptic pA sites.<sup>123</sup> However, U1 snRNP levels in cells can fluctuate, for example as a result of UV-induced DNA damage which causes increased usage of intronic alternative pA sites.<sup>124</sup>

## The polyadenylation step

The second part of the reaction comprises of the non-templated addition of about 200 adenosines,<sup>93,94</sup> which in reconstituted systems only requires the cleaved pre-mRNA template, CPSF, PAP and the poly(A) binding protein (PABPN1, PABPII).<sup>59</sup> PAP is initially tethered to the cleaved pre-mRNA via interactions with the CPSF subunits 160 kDa and Fip1.<sup>29,63</sup> This loose interaction leads to regular dissociation of PAP and adenylation is restricted to intermittent addition of a few adenosines.<sup>59</sup> Subsequently, the nuclear poly(A) binding protein, PABPN1, binds these short oligoadenosine tails added by PAP in its distributive mode.<sup>93</sup> PABPN1-binding provides an additional anchor point for PAP and increases its affinity to the RNA 80-fold.<sup>125</sup> Once this quaternary complex is established on the cleaved transcript, CPSF and PABPN1 cooperatively shift PAP into a processive mode, enabling it to catalyze the full length poly(A) tail without dissociation. The actual number of adenosines added to transcripts is uniform within species, ranging from 70 to 90 in *Saccharomyces cerevisiae*<sup>126</sup> to around 200 in mammals. Extending the poly(A) tail to an appropriate length appears to be critical for appropriate gene expression, as both hypo- and hyperadenylation negatively influence the fate of the mRNA. Hypoadenylated mammalian transcripts have been shown to be retained in the nucleus,<sup>127,128</sup> and inefficient polyadenylation

in yeast is linked to nuclear mRNA degradation.<sup>129</sup> In addition, hyperadenylated host transcripts observed in herpesvirus-infected cells are retained in the nucleus and funnelled into the nuclear mRNA degradation pathway.<sup>130</sup> Furthermore, inhibition of RNA export forces hyperadenylation of nuclear mRNAs.<sup>131</sup>

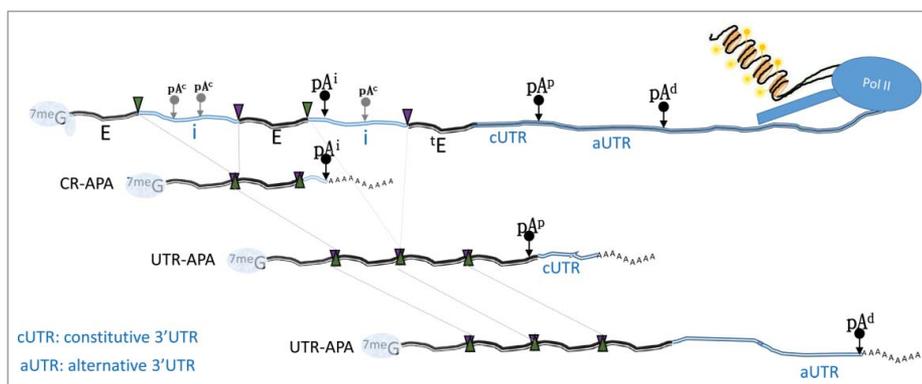
How the precise length of poly(A) tail synthesis is achieved by the polyadenylation complex has been the subject of intensive investigations and much has been learned from reconstituted *in vitro* systems. Central to the model derived from these observations is that the decoration of the poly(A) tail with PABPN1 provides some means of measuring and controlling the number of adenosines added to the 3' end of the cleaved mRNA. The current model envisages that the growing poly(A) tail coated with PABPN1 folds into a spherical complex that promotes contacts between CPSF, PAP and PABPN1. When the tail extends to 250 adenosines, new PABPN1 proteins can no longer be integrated into the sphere and this leads to the disruption of the interactions between CPSF and PAP. Subsequently PAP remains tethered to the RNA only by its interaction with PABPN1 causing it to readopt the distributive mode and ultimately terminate polyadenylation.<sup>94</sup> A direct involvement of PABPN1 in poly(A) length control is further supported by *in vivo* evidence. siRNA mediated depletion of PABPN1 in mouse myoblasts triggered shortening of poly(A) tails and subsequent accumulation of these transcripts in the nucleus.<sup>128</sup> Similarly, nuclear retained transcripts with shortened adenosine tails are characteristic for influenza A virus infected cells. The viral NS1 protein appears to restrict the length of poly(A) tails on host transcripts in infected cells to 12 nucleotides by sequestering PABPN1.<sup>127</sup>

While CPSF, PAP and PABPN1 are sufficient to direct poly(A) tail length control in reconstituted systems, it appears that *in vivo*, more factors are required. The protein nucleophosmin (NPM1) has been found to associate with many mRNAs only after proper termination of polyadenylation.<sup>132</sup> Intriguingly, RNAi knockdown of NPM1 leads to hyperadenylation and nuclear retention of mRNAs suggesting a role of NPM1 in poly(A) tail measurement. The precise molecular mechanisms are not clear but NPM1 somehow contributes to poly(A) tail sizing by regulating the dissociation of the quaternary complex.<sup>132</sup> Perhaps NPM1 functions as a gatekeeper and associates with correctly polyadenylated and export competent transcripts.<sup>133</sup>

## Alternative cleavage and polyadenylation

The complexity of the pA site architecture combined with the multitude of core and auxiliary interacting factors and interconnection with splicing and transcription provides ample opportunities to regulate pA site recognition. It is therefore perhaps not surprising that alternative pA site usage is prevalent in the mammalian transcriptomes with ~70% of all human genes subjected to what is commonly known as alternative cleavage and polyadenylation (APA).<sup>134</sup>

APA occurs when a single gene has multiple pA sites, compared with constitutive polyadenylation, where one gene has one sole pA site. Like alternative splicing, APA results in multiple distinct RNA transcripts being produced from a single gene. The relative positions of these pA sites in a gene and their



**Figure 3.** Coding region APA (CR-APA) and UTR APA. Depending on the location of the different pA sites, APA events can be classed into 2 major groups. CR-APA is the result of differential usage of pA sites that are located within the body of the gene and alternative usage produces APA mRNA isoforms that differ in their coding potential. UTR-APA summarizes events where the different pA sites are located downstream of the stop codon and alternative usage modulates 3'UTR length but does not change the coding potential. pA sites can be found in the intron and in the UTR of a gene. Intronic pA sites (pA<sup>i</sup>) are often cryptic poly A sites (pA<sup>c</sup>) that need to be actively repressed to enable gene expression. pA sites in the 3'UTR are generally separated into proximal (pA<sup>p</sup>) or distal (pA<sup>d</sup>) sites. Usage of the proximal sites generates mRNA isoforms that have a so called constitutive 3'UTR (cUTR) and isoforms that are generated by usage of the distal site contain both the constitutive and alternative 3'UTR (aUTR) regions. The respective resulting APA mRNA isoforms are indicated, dotted lines refer to the removal of introns (i) and fusion of exons (E) and the 5' splice sites and 3' splice sites are indicated by the green and purple triangles respectively. The terminal exon is indicated by "E" and "7meG" refers to the 5' cap.

usage will ultimately determine the coding and regulatory sequence elements that are present in the different transcript isoforms. Depending on where the different pA sites are located in genes, APA events are separated into 2 major categories: untranslated region APA and coding region APA (Fig. 3).

**Untranslated region APA (UTR-APA)** occurs when multiple pA sites are available downstream of the terminal exon of a particular gene. As they are positioned outside the coding regions, differential usage of such pA sites does not alter the coding capacity of the resulting transcript isoforms but instead, alters the lengths of their 3'UTRs. The sequence that is absent in the short isoforms is commonly referred to as the alternative UTR (aUTR), relative to the common UTR (cUTR), which is present in both long and short APA isoforms. The presence or absence of an aUTR may equip a particular mRNA isoform with regulatory elements that influence the fate of the transcript regarding its stability, availability to the translational machinery and subsequent protein output, subcellular localization and even final protein destination. While aUTRs provide scope for potential regulation, it is not always clear what proportion of UTR-APA events have a physiologic impact.

**Coding region APA (CR-APA)** occurs when multiple pA sites are present at positions upstream of the 3'UTR. Thus CR-APA events can produce APA mRNA isoforms that differ in their coding capacity of the final protein product. Cryptic pA sites are also abundant in introns but, as mentioned above, they are actively suppressed by the U1snRNP.

In comparison to UTR-APA, it is easier to delineate a physiologic consequence for CR-APA events as they directly impact on the cellular proteome. Despite this, in the literature global analysis of APA is heavily bias toward UTR-APA.

### Transcriptome wide APA profiles

The sum of all UTR-APA and CR-APA isoforms constitute the cellular APA profile. Advances in next generation sequencing have allowed cellular APA profiles to be analyzed on a transcriptome wide scale and several different pipelines have been developed and are available to investigate APA events (Table 1).

UTR-APA in particular has attracted significant attention in the last decade and a large number of APA profiles have since been determined and are accessible in various repositories including: [<http://www.polyasite.unibas.ch>]<sup>3</sup> [<http://genome.bucm.edu.cn/utr/>].<sup>135</sup> The analysis of these large data sets revealed that APA is highly dynamic and that the profiles can change depending on the context particular cells are studied in.

### APA profiles in the context of proliferating and activated cells

A general pattern of shifts toward the proximal pA sites (UTR shortening) has been seen upon increased cellular proliferation.<sup>145,153</sup> In the case of murine T-lymphocyte activation, the UTR-APA events show a clear pattern of a switch to proximal pA sites, compared with CR-APA, where there appeared to be equal movements toward both proximal and distal pA sites upon cell activation.<sup>153</sup> Comparison with an extensive range of both murine and human cell lines revealed a similar pattern whereby the proliferation rate was inversely correlated with the 3'UTR length.<sup>153</sup>

The activation of neurons by depolarising agents also induces shifts from the distal to proximal pA sites in a subset of genes that are regulated by the MEF2 family of transcription factors.<sup>154</sup> Unlike the examples described above, many of these APA shifts are CR-APA events and thus result in the expression of functionally distinct proteins that may be critical for synapse development.<sup>154</sup> The generality of APA profile shifts in response to cell activation is further shown by shifts that are apparent when astrocytes are exposed to eukaryotic growth factor EGF, T-cells are activated by interleukin-2 and B cells are cultured in the presence of anti-IgM and CpG. Interestingly there appears to be little overlap in genes that are undergoing stimulus dependent APA. This indicates that these shifts are stimulus and also cell type specific,<sup>154</sup> but how they are achieved is unclear in most cases.

Similar to activation, stressed cells tend to enhance the use of intergenic pA sites and produce 3'extended transcripts.<sup>155,156</sup> While no clear trend toward shortening or lengthening in

**Table 1.** Methods of 3' end targeted deep sequencing.

Category	Technique	Overview	References
3' end capture	3P-Seq	A biotinylated primer is added to the poly(A) tail and solely the poly(A) fragments can then be isolated using streptavidin. RT with just TTP is used to fill in poly(A) tail and RNase H is then used to cleave the poly(A) tail leaving just the very 3' end which is then used as input for sequencing library preparation.	136
	3'READS	Uses a unique chimeric CU <sub>5</sub> T <sub>45</sub> oligo isolation system, which completely eradicates internal priming and amplification of oligo adenylated transcripts.	137
Direct RNA sequencing	DRS	DRS uses the Helicos BioSciences system which starts by using an oligo(dT) coated surface to which the poly(A) tail binds to. Reverse transcription using only dTTP is then used to fill in the entire poly(A) tail. Sequencing is then initiated from the most 3' prime non-A base.	28,73,138
Oligo(dT)-based priming	3Seq	Standard oligo(dT)-based priming technique using an oligo(dT) <sub>25</sub> containing RT primer, and sequencing the terminal 25 bp upstream of the cleavage site to map the pA site.	139,140
	3' end RNA-Seq	Standard oligo(dT)-based priming technique using paired-end sequencing to obtain strand-specificity	141,142
	3' end-seq	Includes an <i>in vitro</i> transcription step after reverse transcription to increase the quality of the input RNA and as an amplification step.	143
	3Seq	Adapted slightly from 3Seq	96,144,145
	3'T-fill	Directly before sequencing, the poly(A) tail is filled using TTP resulting in sequencing starting from the base directly upstream of the cleavage site.	146
	A-seq	Uses an RT primer consists of an anchor nucleotide, followed oligo(dT) sequence with a stem-loop containing the adaptor sequence for priming the subsequent PCR reaction, in the middle of the oligo (dT).	147
	MAPS	Standard oligo(dT)-based priming using a random primer for second-strand cDNA synthesis.	148
	PAS-Seq	Variation to the standard oligo(dT)-based priming technique involves using the terminal transferase activity of MMLV reverse transcriptase, which allows generation of cDNAs with linkers in a single RT step, thus skipping several enzymatic steps.	149
	PolyA-seq	Standard oligo(dT)-based priming using a random primer for second-strand cDNA synthesis.	134
	Quant-seq	Commercialised standard oligo(dT)-based priming using a random primer for second-strand cDNA synthesis.	150
SAPAS	Oligo(dT)-based priming using template switching and optimised primer anchoring in the RT-reaction to avoid sequencing in the long poly(A) tail.	135,151,152	
SMPSS	Uses a single-molecule system based on the HeliScope single molecule sequencer. It is amplification and ligation-free, allowing very little bias in quantitation.	97	

UTR-APA has been observed, intronic pA sites are suppressed upon exposure to stress.<sup>157</sup>

### APA profiles in the context of development, tissue and cell types and different species

APA profiles have also been found to vary as a function of the differentiation status of cells. This was first described during mouse embryonic development, where the differentiation of C2C12 myoblast cells to myotubes showed a general shift to distal pA sites (UTR lengthening).<sup>158</sup> Conversely, if differentiation is reversed, and induced pluripotent stem cells are generated from fully differentiated somatic cells, a general 3'UTR shortening is observed.<sup>159</sup> Widespread CR-APA and UTR-APA have been reported to be critical during spermatogenesis. Genes associated with sperm maturation and testis specific genes compared with ubiquitously expressed genes tend to undergo 3'UTR shortening early in spermatogenesis avoiding destabilising and transposable elements that are located in the aUTRs of these genes.<sup>160</sup>

When specific cell types or tissues are compared, they appear to have characteristic APA profiles.<sup>16</sup> The most striking differences in APA profiles are those between mammalian and invertebrate testes and brain. While in testis the proximal pA sites are generally favored, in the brain the opposite is true and a preferred shift toward distal pA sites is characteristic.<sup>161-164</sup> The prevalence of long 3'UTRs in neuronal tissue was reported using northern blot analysis over 20 y ago.<sup>165</sup>

In addition to simple switches between common pA sites, there are several examples of unique pA sites that are only activated in specific tissues such as testis<sup>166</sup> and the brain.<sup>161</sup> In

testis, these sites appear to have a lower incidence of the AAUAAA canonical hexamer, as well as having unique upstream and downstream elements.<sup>166</sup> Since these sites are not used efficiently in somatic cells, it has been suggested that testis-specific pA factors may be responsible for this pA site switching,<sup>167</sup> such as the CstF-64 variant,  $\tau$ CstF-64.<sup>168</sup>

Similarly, tissue specific pA sites are also used in the brain where they generate specific APA mRNA isoforms with unusually long 3'UTRs (> 10 kb).<sup>161</sup> These tissue specific APA events potentially provide a huge repertoire of diverse regulatory elements that can be plugged into tissue specific post-transcriptional regulatory networks.

The conservation of APA sites between species is limited and is restricted to fewer than 500 genes when tandem APA sites are compared between mouse and humans.<sup>169</sup> If human pA sites are compared with rhesus, mouse, rat and dog, fewer than 1% are found in similar regions between them,<sup>134</sup> indicating that APA is chiefly tissue specific, rather than species specific.

Experiments in yeast have compared the conservation of APA between different yeast species including *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, and *Debaryomyces hansenii*.<sup>170</sup> Here, despite the similarities in the general sequence preferences surrounding the pA sites, the APA profiles show species specificity. This led to the conclusion that much of the heterogeneity seen in the pA sites is likely to be biologic noise as defined by "biochemical events that occur *in vivo* as a result of a low-specificity process that has not been subjected to evolutionary optimisation."<sup>170</sup>

Recently, single cell studies have started to expose the heterogeneity of pA site usage within a population of homogenous

stem cells.<sup>171</sup> This revealed that individual cells differ significantly in their choice of pA site usage. Interestingly, it appears that the heterogeneity in pA site choice observed in these cells surpasses that of what would be expected if pA site choice was random, thereby concluding that this variability has the potential to contribute to the functional cell-to-cell heterogeneity.<sup>171</sup>

### APA profiles in the context of cancer

Several studies have observed shifts to usage of proximal pA sites upon cellular transformation and carcinogenesis.<sup>44,138,143,151,153,172-174</sup> It is thought that shifts to these proximal pA sites in cancer function to exclude instability elements present in the aUTR, avoiding particular regulatory networks, such as those involving miRNA-mediated destabilisation of specific transcripts. Accordingly, in many cases shortening can result in increased mRNA and protein levels of oncogenes, for example in *CCND1*.<sup>172</sup> However, it is noteworthy that in the cancer cells that show general 3'UTR shortening, there are also a significant number of genes that undergo 3'UTR lengthening.<sup>138,151,173</sup> This clearly highlights that APA patterns in transformed cells are complex making the interpretation of the profiles in the context of the underlying mechanisms and the physiologic impact of the shifts more difficult.

How cancer specific APA profiles are established is largely unclear, but the dysregulation of CFI in some cancer cells can be linked to 3'UTR shortening and provides a mechanistic explanation for this cancer-specific APA feature.<sup>44</sup>

In addition to further the understanding of how cancer specific gene expression profiles are established, logging changes in APA profiles from tumor cells may also provide a valuable resource that can be exploited to develop tools to type cancer subtypes,<sup>173</sup> provide information for diagnosis,<sup>175,176</sup> and prognosis,<sup>176-178</sup> and perhaps even treatment stratification. Indeed, APA profiles have previously been proposed to serve as biomarkers for cancer progression.<sup>143</sup>

### APA profiles in the context of other disease phenotypes

Many individual APA events have been linked with specific disease phenotypes such as thalassemia and IPEX syndrome.<sup>11,179-181</sup> In these examples, pathogenesis is generally the result of mutations of *cis*-elements in one specific gene that compromises its 3'end processing efficacy. These mutations can result either in loss of function, such as in thalassemia and IPEX syndrome, or gain of function, as exemplified in thrombophilia and mantle cell lymphoma. In the latter, a point mutation in the *CCND1* gene causes activation of a novel pA site leading to 3'UTR shortening and transcript and protein overexpression.<sup>182</sup>

If proteins that regulate cleavage and polyadenylation are mutated, 3'end processing at many different pA sites can be altered and this can lead to changes in APA profiles. For example, the APA profiles between normal and failing hearts differ. While no clear global shifts in 3'UTR lengths are obvious, smaller cohorts of genes in the diseased compared with the normal heart show distinct changes in their APA profiles. Mechanistically the shifts in these APA profiles may be caused by the reduced expression levels of several cleavage and polyadenylation

factors, including PABPN1. Interestingly, the levels of functional PABPN1 have before been linked to global shifts from distal to proximal pA sites in cells from oculopharyngeal muscular dystrophy (OPMD) patients.<sup>96,97</sup> In OPMD cells the availability of functional PABPN1 is compromised by a triplet repeat expansion mutation of an alanine repeat in the PABPN1 gene.<sup>183</sup> It has been suggested that in healthy cells, high levels of PABPN1 are available and can bind and suppress the proximal pA site in a significant number of genes. Conversely, in OPMD cells, the depletion of functional protein results in the activation of proximal pA sites causing 3'UTR shortening which may contribute to the pathology.<sup>96</sup>

Myotonic dystrophy, a neuromuscular disease, represents an additional example where disease specific APA profiles are dependent on the availability of functional proteins. The Muscleblind-like (MBNL) protein family are key regulators of alternative splicing and mutations that compromise their function are associated with myotonic dystrophy. Mutations that cause loss of function in MBNL proteins also result in large scale bidirectional APA shifts.<sup>184</sup>

Extensive APA defects that occur in the cerebellum of amyotrophic lateral sclerosis (ALS) patients have been reported. ALS patients that feature a repeat expansion in the *C9orf72*, have a higher frequency of short 3'UTR APA isoforms. The cause and the physiologic impact of these APA changes however is unknown.<sup>185</sup>

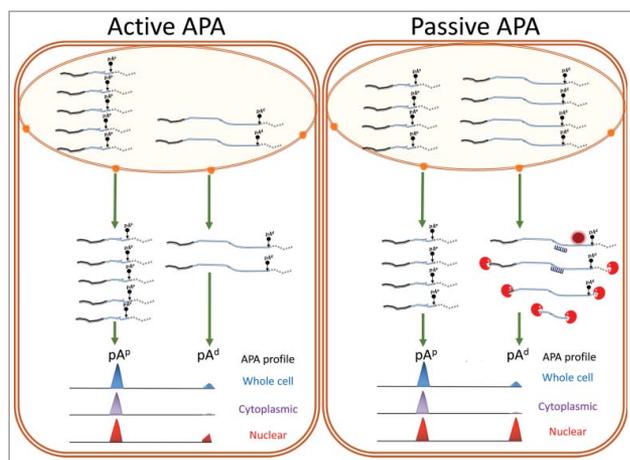
For all these examples it is important to keep in mind that it is not clear whether the shifts in APA profiles are driving disease development or progression, or whether they are a reflection of the disease dependent changes in the cells.

### Issues surrounding the global analysis and interpretation of APA profiles

As outlined above, APA profiles of cells appear to have characteristic signatures that are dynamic and can change in response to specific cues and disease. To elucidate the physiologic role of global APA changes, it is imperative to unravel the mechanisms and networks that are responsible for shaping these profiles. To do this, a distinction has to be drawn between alterations of APA profiles that are due to specific selection of one pA site over another at the point of cleavage, and changes in APA profiles that are due to post-transcriptional mechanisms, such as degradation of specific isoforms.

APA profile changes that occur as a result of cleavage site selection can be regarded as “**active**” APA (Fig. 4, left panel). Active APA can be achieved by changing the availability of factors that either enhance or suppress the usage of a particular pA site over another in response to a specific cue or change in the state of the cell. Conversely, profile changes that are the result of post-transcriptional events can be described as “**passive**” APA and may occur in response to a specific cue that alters the availability of factors, for example miRNA levels, that act post-transcriptionally (Fig. 4, right panel). This includes processes that alter the APA profile by interfering with the nuclear export rates or the stability of specific APA mRNA isoforms.

As most APA analysis is performed using whole-cell RNA, active and passive APA events are difficult to distinguish (Fig. 4). This may complicate the evaluation of particular



**Figure 4.** Distinguishing between ACTIVE and PASSIVE APA. The APA profile can either be modified at the point of cleavage (Active), or at the post transcriptional level (Passive). In active APA, factors that inhibit or enhance one pA site over another produce APA isoforms that can avoid a particular regulatory pathway. On the other hand, in passive APA, the availability of factors such as RBPs (dark red circle) and miRNAs (navy) in the cytoplasm alter the APA profile by specifically down-regulating a particular isoform. For example, as depicted here, miRNAs can target the aUTR which can recruit the RNA induced silencing complex (RISC) result in degradation by exoribonucleases (red "PacMan"). Different RBPs that bind to the aUTR can either stabilize or degrade the isoform. In this case although the whole cell APA profile is the same, the nuclear APA profile is different, highlighting the importance of assessing changes in the cytoplasm compared with the nucleus to distinguish Active and Passive APA. This gives a better resolution of the causes that enforce specific APA changes in different environments.

profiles and may lead to misinterpretations of the underlying mechanisms and thus the physiologic relevance of APA. For example, shortening events in the context of cancer are often attributed to the selection of proximal sites at the point of cleavage to avoid post-transcriptional control executed by miRNAs. This interpretation however overlooks the fact that miRNA levels are known to be subjected to dramatic changes during oncogenic transformation<sup>186</sup> and that different cancer subtypes have distinctive miRNA signatures.<sup>187</sup> Thus a significant number of the shortening or lengthening events may be due to passive APA acting at the post-transcriptional level. Liaw *et al.* highlight this possibility by cross-referencing miRNA profiles with previous APA data showing apparent 3'UTR shortening in the breast cancer cell line, MCF-7.<sup>151,188</sup> This analysis showed that genes with target sites for upregulated miRNAs in their aUTR exhibited a larger degree of 3'UTR shortening relative to those without targets, and the extent of this was influenced by the number of target sites. In addition, the positioning of target sites for miRNAs upregulated in MCF-7 cells showed that they were enriched in the aUTRs of genes where 3'UTR shortening was seen, relative to downregulated miRNAs. Combining this data, it was concluded that selective degradation of solely the longer transcript isoforms by upregulated miRNAs has a significant impact that contributed to the observed changes in APA profiles in MCF-7 cells.<sup>188</sup>

Perhaps many of these passive APA events in cancer cells are "collateral damage" caused by the upregulation or downregulation of particular miRNAs that aim to modulate a handful of specific genes. However, due to the nature of miRNA-mediated regulation, this will also affect many other genes that have

little, if any, consequence for the progression or establishment of tumorigenesis.

The most common method used to untangle active and passive APA events is to look at the expression levels of all transcript isoforms.<sup>144,153,177</sup> This approach argues that if observed shifts in APA patterns are due to a change in pA site choice, the increased proportion of the shorter transcripts leads to a lower overall decay rate, as 3'UTRs overall are of a repressive nature.

The expression level observed would, therefore, increase in this situation. Conversely, if the shift in APA pattern is due to an increase in the decay rate of solely the longer transcript, a decrease in expression would be seen. This method has its limitations, as any number of mechanisms, aside from APA, can influence transcript levels of a gene. To address this issue more directly, APA profiling of different subcellular fractionations have been used.<sup>16,189</sup> In particular, the comparison between nuclear and cytoplasmic APA events can be used to get a better resolution on specific APA events and determine whether they are due to a change in cleavage site choice, or due to regulation at the post-transcriptional level including nuclear retention and transcript stability<sup>16</sup> (Fig. 4). If subcellular APA is performed between different states of cells it may reveal how APA events are integrated with other regulatory networks that dynamically respond to changes in the state of cells.

### Factors that regulate active APA by influencing pA sites choice at the point of cleavage

Despite the large number of reported shifts between pA sites in different cells and cell states, the precise mechanism of defining pA site choice in active APA remains ill defined. This is partly due to the complexity by which the strength of a pA site is defined and the many ways this can be modulated. In addition, as mentioned above, the difficulty to discriminate between active and passive APA further complicates the identification of factors. As the *cis*-elements influencing polyadenylation efficiency do not change, the switching between pA sites can only be achieved through the fluctuation of *trans*-factors or the presentation of pA sites to the processing machinery during transcription. The former is exemplified by the transplant of large chromosomal segments from *Debaryomyces hansenii* into *Saccharomyces cerevisiae* and the concomitant adoption of the host APA profile by the recipient.<sup>170</sup> Several *trans*-factors have been shown to influence pA site choice (Fig. 5), including core pA factors (Table 2) and several RNA-binding proteins of diverse function (Table 3).

As can be seen from Tables 2, 3 and 4, there is a plethora of *trans*-factors that, when depleted or compromised, cause shifts in APA profiles.

A number of RNA binding proteins have been associated with influencing pA site usage, that primarily have other nuclear functions, as outlined in Table 3. Many of these mediate splicing, again highlighting the interconnection between these 2 pre-mRNA processing mechanisms.

As summarised in Table 4, several features that modulate gene transcription have also been linked to active APA. It has been long known that pre-mRNA processing and



**Table 3.** Other RNA-binding factors known to influence pA efficiency.

Factor	Motif bound	Proposed model	References
$\alpha$ CP ( $\alpha$ CP)	C-rich motifs	$\alpha$ CP binds mRNAs containing a subgroup of C-rich elements in their UTRs and acts as an upstream 3' end processing enhancer. Usage of distal or proximal pA sites can be influenced depending on upstream C-rich regions close to the respective pA site by varying $\alpha$ CP levels.	198
Cirbp and Rbm3	GNNGNNG	Upon cold-shock, these factors are upregulated and, through 3' UTR binding, inhibit the use of proximal pA sites.	199
CPEB1	CPE	CPEB1 shuttles to the nucleus binding cytoplasmic polyadenylation elements and enhances polyadenylation at nearby pA sites. Also, it prevents U2AF65 binding, which inhibits splicing.	200
DICER		CPEB1 in the nucleus causes shortening and this correlates with cell proliferation.	16
ELAV ( <i>Drosophila</i> )		Nuclear Dicer affects pA site usage by modifying the chromatin landscape surrounding the 3' end processing sites. In a region of closed chromatin Pol II progression is slowed down, increasing the likelihood that a weak pA site is recognized. In contrast if the weak pA site is in an open conformation, Pol II progression is fast decreasing pA site usage.	201,202
FUS	UGGUU	In the neuronal tissues, ELAV is recruited to the promoter-paused Pol II complex. Upon resuming transcription, ELAV is deposited near proximal pA sites, inhibiting their usage, resulting in extended 3' UTRs.	203
hnRNP C	U-rich	FUS binds directly downstream of a proximal pA site, which enhances CPSF160 recruitment and activates the pA site leading to short transcripts. If there is no pA site upstream of a FUS binding site, FUS binding causes Pol II stalling and premature termination, producing short transcripts that are not polyadenylated.	3
hnRNP F	G-rich DSE	hnRNP C binds to U-rich sequences, which masks the pA site in its vicinity to represses their use. The transcripts affected by hnRNP C mediated APA are enriched in ELAVL1 binding sites and this process may thus be linked to the HuR ( <i>ELAVL1</i> ) mediated recruitment of specific mRNA isoforms to the Endoplasmic Reticulum.	204
hnRNP H1	Auxiliary DSE	Competes with CstF-64 by binding to G-rich motifs near pA sites.	205
hnRNP H2	G-rich	Depletion results in a general shift to distal pA sites, with hnRNP H1 binding sites surrounding proximal pA sites.	206
hnRNP K	UCCCUU	Binds near pA sites and enhances binding of CstF-64.	207
hnRNP L	CA-rich elements	Competes with CFI for binding the pre-mRNA, reducing pA efficiency and reduced usage of that pA site.	208
HuR ( <i>ELAVL1</i> )	AU-rich elements (AREs)	Functions as a splicing regulator, so altering levels of hnRNP L can sway the balance between competing splicing and intronic pA events	209,210,211
Mbnl proteins	R/YGCV	HuR ( <i>ELAVL1</i> ) binds to AREs in transcripts. Through competition for binding sites on pre-mRNAs, HuR can influence both APA and splicing, including its autoregulation. When associated with particular aUTRs, HuR can also control the final destination of the protein product. For example, the <i>CD47</i> aUTR mRNA isoform protein product is preferentially located to the plasma membrane, whereas the short UTR isoform lacking HuR sites translates <i>CD47</i> that remains predominantly in the ER.	184
MED23		Muscleblind-like proteins (Mbnl) are important regulators of alternative splicing during development. Mbnl is also implicated in APA and can either inhibit pA site usage if it binds close to a pA site or enhance pA site usage if it binds further upstream. Inhibition is thought to occur through steric hindrance. Mbnl is critical for creating a normal APA landscape during development and dysregulation of this process is associated with myotonic dystrophy.	212
Nkx2-5		Mediator complex subunit 23 (MED23) interacts with hnRNP L and affects hnRNP L regulated APA events, possibly by controlling hnRNP L occupancy at the promoter.	213
Nova	YCAV	In conjunction with Xrn2, Nkx2-5 regulates pA site usage which is of high importance during mouse heart development. This tissue specifically expressed factor regulates APA, and its knockdown causes 3' UTR lengthening.	214
Paf1C		NOVA is a neural-specific factor that binds YCAV elements in the 3' UTR. Depending on the location of these motifs, binding of NOVA can influence pA site choice by suppressing their use.	215
PTB	G-rich USE	Depletion of some Paf1C subunits (Paf1, Cdc73, Ski8) results in global 3' UTR shortening. Regarding CR-APA, only Paf1 and Cdc73 depletion activated coding region pA sites. Paf1C subunits also play a role in suppressing transcription site intronic pA sites. Absence of Paf1 may cause increased Pol II pausing, which stimulates recognition of a pA site in the coding region.	216
RBBP6	unknown	PTB competes with CstF64 to bind the DSE, thereby inhibiting pA site usage. However, it can also aid recruitment of hnRNP H1, which stimulates pA site usage.	217
SRm160	Unknown	RBBP6 competes with its isoform iso3 for binding with the core pA machinery. When RBBP6 is bound, it enhances pA site cleavage efficiency and promotes the use of weaker proximal pA sites. RBBP6 and iso3 particularly affect APA in transcripts that have AU-rich 3' UTRs such as c-jun.	218
SRSF3	CNUC	Enhances pA through the association with CPSF.	219
SRSF7		Promotes biogenesis of long 3' UTR APA isoforms and regulates their nuclear cytoplasmic export.	219
TDP-43	UG rich	Promotes biogenesis of short 3' UTR APA isoforms and regulates their nuclear cytoplasmic export.	220
THOC5		High levels of TDP-43 cause inhibition of pA1 site in intron 7 of its own <i>TARDBP</i> pre-mRNA, resulting in usage of pA2 or pA4, which both produce transcripts that are targeted by the Nonsense Mediated Decay pathway, providing a mechanism of auto-regulation.	65
U1 snRNP	AGGURAGU	THOC5 is a member of the human transcription export complex (TREX). THOC5 knockdown activates proximal pA site usage. It is suggested that THOC5 recruits CF <sub>m</sub> 68 to target genes, promoting distal pA site usage.	121,122,124,221,222
		Suppresses cryptic pA sites in the gene body, which is essential for the formation of full-length transcripts. Shown to suppress premature transcription termination in polycistronic pre-mRNAs in <i>C. elegans</i> by inhibition of pA site recognition by Pol II associated poly(A) factors. U1 snRNA levels drop after UV-induced DNA damage and activate intronic pA sites.	

(Continued on next page)

Table 3. (Continued)

Factor	Motif bound	Proposed model	References
U1A	AUGCN(1–3)C	Component of U1 snRNP which binds GU-rich regions downstream of pA sites inhibiting the binding of CstF64, thus inhibiting polyadenylation. U1A can also bind to PAP inhibiting the polyadenylation reaction itself. U1A is known to inhibit polyadenylation of its own mRNA, and has also been shown to act independently of U1 snRNP to inhibit polyadenylation of the <i>SMN</i> transcript.	223,224
U2	3SS	U2 interacts with CPSF and enhances polyadenylation efficiency.	114
U2AF	Pyrimidine tract	U2AF interacts with CFI stimulating pA.	111

slowing of the polymerase which would favor the more proximal positioned sites.<sup>215</sup> Furthermore, the promoters of specific genes in *Drosophila*, have been shown to be essential in the recruitment of ELAV to Pol II, which subsequently acts at the 3' ends of those genes to inhibit proximal pA site, resulting in 3'UTR extension.<sup>201</sup> The upregulation of specific transcription factors in response to cell activation, for example E2F, can influence APA profiles indirectly by increasing the gene expression of key 3' end processing factors.<sup>145</sup>

### The impact of active APA on the regulation of gene expression

Modulating the length of the UTRs by APA has the potential to produce mRNA isoforms that can be subjected to distinct post-transcriptional regulation (Fig. 6). As 3'UTRs are thought to have an overall repressive role,<sup>144</sup> cleavage and polyadenylation of a transcript at a pA site more proximal to the stop codon is generally considered to result in the removal of potentially repressive *cis*-elements in the 3'UTR, summarised in Table 5. This can, therefore, stabilize the transcript, or promote its translation, thus increasing the overall expression of that gene (Fig. 6). However, several RNA binding proteins have been identified to bind to the 3'UTR and stabilize the transcript, for example the Hu family of proteins (including HuR), which bind to AREs<sup>236,237</sup>. The neuronal specific HuD protein stabilizes transcripts required in neurite outgrowth enabling neuronal differentiation in several cell lines.<sup>238,239</sup>

### The impact of APA on transcript isoform stability and translation

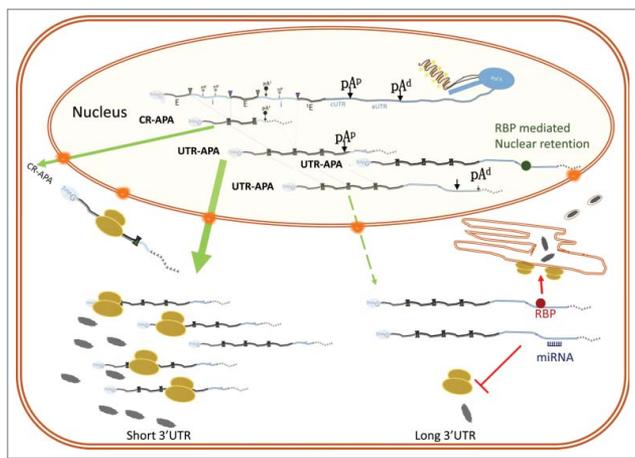
Global comparisons of *cis*-elements located in the aUTRs compared with cUTRs revealed a bias toward a higher representation of conserved miRNA seed regions in aUTRs.<sup>153,158,248</sup> The miRNA target sites in aUTRs are also located in regions with higher AU content. This reduces the possibility that these target sites are involved in secondary structures and thus would represent better targets for miRNAs.<sup>158</sup>

Individual gene examples such as *PAX3*, *CCND1*, *CCND2*, *CDC6*, *MAPT* and *MGMT* support that the biased distribution of miRNA binding sites in APA mRNA isoforms can act as a potent layer to govern miRNA mediated regulation of gene expression.<sup>172,249–252</sup> Similarly, APA in *ELAVL1* and *PDCL* regulates their expression by controlling the presence or absence of destabilising or stabilizing RNA binding protein recognition motifs (Table 6).<sup>253,254</sup>

However, despite these individual gene examples and the general biased distribution of *cis*-elements, the actual global impact of APA on isoform stability and transcript levels has been found to be fairly modest.<sup>97,144,145,151</sup> This was further supported by more detailed analyses addressing the impact of UTR shortening on mRNA isoform half-life times in proliferating T-cells<sup>265</sup> and mouse 3T3 cells, where aUTRs showed a limited influence on the stability of APA mRNA isoforms.<sup>257</sup> In addition, contrary to expectations, a significant proportion of APA mRNA isoforms with extended UTRs have been

Table 4. Features and conditions that can influence pA site choice.

Factor	Motif bound or affected	Proposed model	References
DNA methylation (imprinting)	CpG islands	The methylation status of CpG islands influences pA site selection in the murine imprinted gene <i>H13</i> . This is thought to act via an unidentified DNA methylation sensitive pA factor. Similarly, in the <i>Herc3/Nap1i5</i> locus, the regulation of an internal pA site and the downstream pA site that produces full length <i>Herc3</i> mRNAs is controlled by DNA methylation at internal CpG islands.	234,235
E2F	TTGGCGG	Through enhanced proliferation, increased levels of the transcription factor E2F result in the increased use of proximal pA sites by upregulation of key 3' end processing genes.	145
Nucleosome positioning		High nucleosome occupancy directly upstream of proximal pA sites generally correlates with increased proximal pA usage.	231,232
Transcription rate		Slow transcription rates result in a longer time between when the proximal and distal pA sites are transcribed thereby causing in an increased probability of proximal pA site utilization. At the single gene level, pausing downstream of the intronic ( $\mu$ S) pA site in the <i>IGHM</i> gene can coordinate 3' end processing factor recruitment and pA site usage.	16,230,234
H3K4me3 levels		Chromatin status regulates pA site choice. An "open chromatin" state as measured by high H3K4me3 levels in spermatids compared with spermatocytes influences pA site usage resulting in global UTR shortening accompanied with greater transcript stability.	160
Neuronal activity		Neuronal activity promotes the use of proximal and internal pA sites affecting many transcription factor MEF2 target genes.	154
Stress: arsenite, anisomycin, viral stress		Viral stress or cells exposed to stress agents such as arsenite and anisomycin tend to enhance the usage of intergenic pA sites and generate 3' extended transcripts. Anisomycin mediated stress also suppressed intronic pA sites and pA sites that are located in the ORF. No clear trend is observed regarding 3'UTR-APA events.	155,157



**Figure 6.** Consequences of APA: APA-isoform dependent decay rates and protein output. The 3'UTR length changes arising from APA can have implications on mRNA localization and transcript stability, which can impact on protein output and also determine the final destination of the encoded protein. This figure depicts the case where a short 3'UTR evades miRNA target sites in the aUTR, making it a more stable transcript, enabling increased protein output (protein symbolised by gray globules; ribosomes symbolised by mustard colored structures). The longer isoform shown here is bound by an RBP (dark green) in the nucleus, which prevents its export into the cytoplasm. The transcripts that are exported can be targeted for degradation by miRNA binding to the aUTR. The aUTR of the longer isoform can also bound by an RBP (dark red circle) in the cytoplasm which alters the localization of the transcript, for example in close proximity to the Endoplasmic Reticulum, for protein synthesis. Therefore, the UTR is important in mediating nuclear export, transcript stability, translatability and mRNA localization and the modulation of this is achieved by changing the expression of RBPs and miRNAs.

shown to have increased stability in 3T3 cells,<sup>257</sup> a phenomenon that was also observed in adipocyte stem cells.<sup>258</sup> Following this trend, a recent analysis of nuclear and cytoplasmic fractions confirmed that the impact of UTR-APA on the stability is modest.<sup>16,259</sup> In HEK293 cells, around 10 percent of all APA events are found to be subject to post-transcriptional regulation and around 3% of all those cytoplasmic events are destabilised by miRNAs. Interestingly, this study further showed that miRNAs target both short and long UTR-APA isoforms for degradation.<sup>16</sup>

The lack of a global impact of APA on mRNA stability in mammalian cells however is not universal. Contrary to mammals, in yeast, changes in UTR length have comprehensive implications on transcript stability and even changes at

a single nucleotide level can have a significant large scale impact on transcript stability.<sup>260,261</sup>

The impact of different APA mRNA isoforms on translation efficiency follows a similar pattern. While the APA mediated shortening of UTRs in several mRNAs including *CDC6*,<sup>251</sup> *HSPA2*,<sup>262</sup> *ECE-1*,<sup>263</sup> *CCND2* and *DICER1*<sup>174</sup> was concomitant with a higher translational output, the opposite was true for many genes in 3T3 cells.<sup>257</sup> Most notably in the *SERT* gene (serotonin receptor), the distal isoform is translated at a higher rate.<sup>264</sup> As for the impact on stability, APA controlled UTR shortening appears to have a limited global impact on translation efficiency.<sup>256,257</sup>

While there is not a clear global trend, it is nevertheless clear from these examples that the differential usage of pA sites can alter the stability of selected transcripts and affect their overall protein expression, consequently having a significant impact on major biochemical processes.

### UTR-APA and the control of subcellular localization of transcripts

UTRs can play an important role in directing the subcellular localization of transcripts<sup>265</sup> and it is therefore not surprising that UTR-APA influences subcellular distribution of mRNA isoforms. 3'UTR lengthening in the brain is one of the most striking examples of tissue-specific APA<sup>161</sup> and this also appears to provide an important platform to control the localization of APA mRNA isoforms. *BDNF*,<sup>266</sup> *Ranbp1*,<sup>267</sup> *Impa1*,<sup>268</sup> *MKK7*<sup>269</sup> and *KPNB1*<sup>270</sup> are examples where APA isoforms with long 3'UTRs include localization signals resulting in targeted transport and translation of the respective transcripts. Similarly, 3'UTR extension in  $\alpha$ -synuclein transcripts, which are associated with Parkinson disease pathology, not only increases the translational output but also affects their localization away from the synaptic-terminals to mitochondria.<sup>271</sup> On a more global scale, a recent comparison between RNA isolated from the dendrites and the stroma in neurons found UTR-APA isoforms that display differential subcellular localization. Distal APA isoforms that are induced during neuronal differentiation are more likely to be localized in neurite projections than their proximal APA counterparts.<sup>189</sup>

APA mediated subcellular localization is not just restricted to highly polarized cells. At the individual gene

**Table 5.** Cis-elements in the 3'UTR.

class of cis-elements	Sequence element	Overview	References
AU-rich elements (AREs)	AUUUA	These are present in 5–8% of all genes and can trigger mRNA destabilisation and translational repression. This is triggered by the binding of ARE-binding proteins (ARE-BPs), including TTP. The Hu family of proteins bind AREs and stabilize the corresponding transcript, particularly during neuronal differentiation.	240–242
GU-rich elements (GREs)	GUUUG	Contained in at least 5% of human mRNAs and triggers mRNA deadenylation and degradation. Acts through binding of proteins from the CELF family.	243
CU-rich elements (CUREs)	(C/U)CCAN <sub>x</sub> CCC	PTB is the best-characterized CURE-binding protein and can affect translational repression, polyadenylation and mRNA stability.	244
CA-rich elements (CAREs)	(U/A)Py <sub>x</sub> UC(C/U)CC (CA) <sub>n</sub>	A stabilizing dinucleotide repeat, which acts primarily via hnRNP L binding, which alters the susceptibility of the mRNA to endo- and exonucleases.	245
microRNA target sites	NNNNNNN	By far the most common destabilising element and target sites are present in > 60% of all genes. Regulation is primarily done via destabilisation of target mRNA (> 84%), rather than translational inhibition.	246,247

**Table 6.** Examples of genes producing differentially regulated UTR-APA isoforms.

Gene	Gene function	Summary	References
<i>CCND2</i>	Cell cycle regulator	3'UTR shortening is seen in cancer cell lines relative to normal tissues, thereby avoiding regulation by miR-15/16. Preferential use of the proximal pA site has been shown to increase the number of cells present in S-phase.	172
<i>CDC6</i>	Cell cycle regulator	Usage of the proximal pA sites avoids miRNA-mediated repression, resulting in increased CDC6 protein levels. This is triggered by the potent proliferation signal 17 $\beta$ -estradiol (E2), and may, therefore, be a mechanism by which the cell promotes cell cycle progression in response to proliferation signals.	251
<i>ELAVL1</i> (HuR)	RNA-binding protein	The aUTR region of HuR mRNA region contains an ARE region where HuR and TTP competitively bind, resulting in mRNA stabilization or destabilisation respectively. This, therefore, creates an autoregulatory loop, which may amplify the pathological role of HuR.	254
<i>MAPT</i>	Stabilises microtubules, specifically in neurons	UTR-APA isoforms are differentially regulated in neuroblastoma cell lines, with miR-34 family members targeting solely the distal APA isoforms. This gene encodes the Tau protein, which is one of the key components of protein aggregates formed during Alzheimer disease.	255
<i>MGMT</i>	DNA repair	Glioblastomas are shown to shift pA site usage to a distal site, resulting in the inclusion of target sites for miR-767-3p, miR-181d and miR-648, thus reducing the expression of <i>MGMT</i> .	250
<i>PAX3</i>	Transcription factor which controls myogenesis	In quiescent muscle stem cells, APA results in the production of <i>PAX3</i> , a key myogenic regulator, with a shortened 3'UTR allowing escape from regulation via miR-206 that targets the aUTR.	249
<i>PDCL</i>	Putative modulator of heterotrimeric G proteins	Several AREs are located between 2 pA sites in the 3'UTR of <i>PDCL</i> , which results in a significantly shorter mRNA half-life of the longer transcript.	253
<i>ZFR</i>	Neuron development	<i>ZFR</i> APA isoforms are differentially regulated by miR-579, which itself is co-transcribed with the <i>ZFR</i> gene. miR-579 also regulates <i>CPSF2</i> , creating a negative feedback loop wherein transcription of <i>ZFR</i> results in the production of miR-579, which targets <i>CPSF2</i> , favoring usage of the proximal pA site, which is resistant to regulation by miR-579. Transcription of <i>ZFR</i> to produce miR-579 also regulates <i>CPSF2</i> in a negative feedback loop. The longer <i>CPSF2</i> isoform is targeted by miR-579, favoring the usage of the proximal pA site, which is resistant to regulation by miR-579.	252

level, it has been shown that the presence of *Alu* repeats in the aUTRs of the *Nicolin1* gene is incompatible with nuclear cytoplasmic export in HEK293 cells.<sup>272</sup> In a more recent global study, a significant number of nuclear retained APA mRNA isoforms have been identified when nuclear and cytoplasmic APA profiles in HEK293 cells were compared. At least for some of these APA isoforms incomplete splicing resulting in intron retention is instrumental to prevent export into the cytoplasm.<sup>16</sup> In addition, APA isoforms with short 3'UTRs tend to be overrepresented in the cytoplasm compared with the nucleus and while this phenomenon is conserved between different cell types, the genes affected are cell type specific.<sup>16</sup> Interestingly, 2 prominent splicing factors, SRSF3 and SRSF7 have recently been linked to both pA site selection and selective nuclear cytoplasmic transport. SRSF3 promotes the usage of distal pA sites and by recruiting the export factor NXF1, controls the export of long 3'UTR APA isoforms. In contrast, SRSF7 causes a shift toward proximal pA site usage and subsequently supports the export of the resulting short UTR-APA isoforms.<sup>219</sup>

UTR-APA controlling localized translation has also been seen with some membrane bound proteins including *CD47*. The distal APA isoform of *CD47* allows binding of HuR and SET resulting in the translated protein to be localized to the plasma membrane. Conversely, the proximal APA isoform translated *CD47* locates primarily to the endoplasmic reticulum.<sup>211</sup>

### The impact of CR-APA

In contrast to UTR-APA, CR-APA by definition, will affect the protein identity, as introducing a pA site upstream of a translational stop codon will result in a protein with a different C-terminus. This may be a simple truncation in the case of exonic

CR-APA, or more commonly a different C-terminus in composite-intronic CR-APA or skipped-intronic CR-APA (Fig. 1). In the case of skipped-intronic APA, polyadenylation is coupled to an alternative splicing event of the alternative terminal intron and is therefore in dynamic competition with splicing.<sup>273</sup> Thousands of these intronic pA sites have been identified, but largely remain dormant through inhibition by factors such as U1 snRNP<sup>73</sup> (see above). However, these intronic pA sites can indeed be activated, as is seen during increased proliferation.<sup>145</sup> In around 7.5% of breast cancer patients, a short version of *MAGI3* as a result of premature polyadenylation at a cryptic pA site located in intron 10 of the gene is present. The shortened MAHI3 protein product promotes mammary cell transformation. However, the activation of this cryptic pA site is unclear as it is not dependent on mutations of local cis-elements and thus is more likely to involve an imbalance of trans-factors.<sup>156</sup>

The classic example of physiologic relevant CR-APA is that of the switch between the membrane-bound form of IgM in B-cells to the secreted form in plasma cells.<sup>274</sup> This switch is thought to be governed by changes in the CstF-64 levels, whereby an elevation of CstF-64 in plasma cells activates the weaker proximal pA site, triggering the switch to the upstream pA site, ultimately removing the membrane anchoring C-terminus.<sup>193</sup>

One important constraint of CR-APA isoforms is that they do not result in the introduction of premature translational termination codons. Therefore, they are not subject to nonsense-mediated decay and have the capability of producing alternative protein isoforms with potentially distinct physiologic roles.<sup>275</sup> CR-APA events also include examples where the polyadenylation event within the coding region of an exon can convert a tyrosine codon (TAT) into a stop codon (TAA). This mechanism regulates expression of a truncated form of glutamyl-prolyl tRNA synthetase that enables

the escape of inflammatory genes from the GAIT complex mediated translational silencing.<sup>276</sup>

## Outlook

The advent of high-throughput sequencing technologies has no doubt made a huge contribution to our understanding of 3' end formation and APA. It has allowed us to map pA sites on a transcriptome wide scale and compare their differential usage in tissues and diverse cellular states. While in the past the focus was on mapping "global" trends such as 3'UTR lengthening and shortening in different cellular contexts, in the future the focus has to shift more toward identifying APA events that have a proven physiologic impact. This requires a clear distinction to be made between active and passive APA events. Furthermore, rather than just bioinformatics correlations, the physiologic consequence of such APA events also has to be confirmed by direct experimentation at the individual gene level.

The appreciation of APA as a global regulator of gene expression in recent years has also revived a renewed focus on the mechanistic side of pA site recognition. Despite the simplicity of the reaction, it is still unclear how the pA machinery can assemble into functional complexes at pA signals that deviate considerably from the consensus sequences and differ widely in their architecture. Developing methods that can assess the likelihood of a pA site being subjected to regulation<sup>277</sup> and further dissecting the mechanistic details of pA site recognition will also be critical to fully understand APA and its physiologic impact.

## Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

## Acknowledgments

We thank Nick Proudfoot for the critical reading of the manuscript. AF's research is funded by the BBSRC (BB/N001184/1) and JN and RP are funded by the MRC.

## ORCID

Radhika Patel  <http://orcid.org/0000-0002-6880-0149>  
 Zhiqiao Wang  <http://orcid.org/0000-0002-3453-3855>  
 Alastair Louey  <http://orcid.org/0000-0003-4201-4729>  
 André Martin Furger  <http://orcid.org/0000-0002-1161-2841>

## References

1. Evguenieva-Hackenberg E, Klug G. New aspects of RNA processing in prokaryotes. *Curr Opin Microbiol* 2011; 14:587-92; PMID:21945217; <https://doi.org/10.1016/j.mib.2011.07.025>
2. Proudfoot NJ, Furger A, Dye MJ. Integrating mRNA processing with transcription. *Cell* 2002; 108:501-12; [https://doi.org/10.1016/S0092-8674\(02\)00617-7](https://doi.org/10.1016/S0092-8674(02)00617-7)
3. Gruber AJ, Schmidt R, Gruber AR, Martin G, Ghosh S, Belmadani M, Keller W, Zavolan M. 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; 26:1145-59; PMID:27382025; <https://doi.org/10.1101/gr.202432.115>
4. Beaulieu E, Freier S, Wyatt JR, Claverie JM, Gautheret D. Patterns of variant polyadenylation signal usage in human genes. *Genome Res* 2000; 10:1001-10; PMID:10899149; <https://doi.org/10.1101/gr.10.7.1001>
5. Tian B, Hu J, Zhang H, Lutz CS. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res* 2005; 33:201-12; PMID:15647503; <https://doi.org/10.1093/nar/gki158>
6. Zarudnaya MI, Kolomiets IM, Potyahaylo AL, Hovorun DM. Downstream elements of mammalian pre-mRNA polyadenylation signals: Primary, secondary and higher-order structures. *Nucleic Acids Res* 2003; 31:1375-86; PMID:12595544; <https://doi.org/10.1093/nar/gkg241>
7. Hu J, Lutz CS, Wilusz J, Tian B. Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. *RNA* 2005; 11:1485-93; PMID:16131587; <https://doi.org/10.1261/rna.2107305>
8. Fitzgerald M, Shenk T. The sequence 5'-AAUAAA-3' forms parts of the recognition site for polyadenylation of late SV40 mRNAs. *Cell* 1981; 24:251-60; PMID:6113054; [https://doi.org/10.1016/0092-8674\(81\)90521-3](https://doi.org/10.1016/0092-8674(81)90521-3)
9. Montell C, Fisher EF, Caruthers MH, Berk AJ. Inhibition of RNA cleavage but not polyadenylation by a point mutation in mRNA 3' consensus sequence AAUAAA. *Nature* 1983; 305:600-5; PMID:6194440; <https://doi.org/10.1038/305600a0>
10. Wickens M, Stephenson P. Role of the conserved AAUAAA sequence: Four AAUAAA point mutants prevent messenger RNA 3' end formation. *Science* (80-) 1984; 226:1045-51; PMID:6208611; <https://doi.org/10.1126/science.6208611>
11. Orkin SH, Cheng TC, Antonarakis SE, Kazazian HH. Thalassemia due to a mutation in the cleavage-polyadenylation signal of the human beta-globin gene. *EMBO J* 1985; 4:453-6; PMID:4018033
12. Chen J-M, Férec C, Cooper D. A systematic analysis of disease-associated variants in the 3' regulatory regions of human protein-coding genes I: General principles and overview. *Hum Genet* 2006; 120:1-21; PMID:16645853; <https://doi.org/10.1007/s00439-006-0180-7>
13. Danckwardt S, Hentze MW, Kulozik AE. 3' end mRNA processing: Molecular mechanisms and implications for health and disease. *Embo J* 2008; 27:482-98; PMID:18256699; <https://doi.org/10.1038/sj.emboj.7601932>. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=18256699](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18256699)
14. Higgs DR, Goodbourn SE, Lamb J, Clegg JB, Weatherall DJ, Proudfoot NJ. Alpha-thalassaemia caused by a polyadenylation signal mutation. *Nature* 1983; 306:398-400; PMID:6646217. Available from <http://www.ncbi.nlm.nih.gov/pubmed/6646217>
15. Ferreira PG, Oti M, Barann M, Wieland T, Ezquina S, Friedländer MR, Rivas MA, Esteve-Codina A, Rosenstiel P, Strom TM, et al. Sequence variation between 462 human individuals fine-tunes functional sites of RNA processing. *Sci Rep* 2016; 6:32406; PMID:27617755; <https://doi.org/10.1038/srep32406>
16. Neve J, Burger K, Li W, Hoque M, Patel R, Tian B, Gullerova M, Furger A. Subcellular RNA profiling links splicing and nuclear DICER1 to alternative cleavage and polyadenylation. *Genome Res* 2016; 26:24-35; PMID:26546131; <https://doi.org/10.1101/gr.193995.115>. Available from <http://genome.cshlp.org/lookup/doi/10.1101/gr.193995.115>
17. Zhao J, Hyman L, Moore C. Formation of mRNA 3' ends in eukaryotes: Mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev* 1999; 63:405-45; PMID:10357856. Available from <http://www.ncbi.nlm.nih.gov/pubmed/10357856>
18. Nunes NM, Li W, Tian B, Furger A. A functional human poly(A) site requires only a potent DSE and an A-rich upstream sequence. *EMBO J* 2010; 29:1523-36; PMID:20339349; <https://doi.org/10.1038/emboj.2010.42>
19. Danckwardt S, Gehring NH, Neu-Yilik G, Hundsdoerfer P, Pforstich M, Frede U, Hentze MW, Kulozik AE. The prothrombin 3' end formation signal reveals a unique architecture that is

- sensitive to thrombophilic gain-of-function mutations. *Blood* 2004; 104:428-35; PMID:15059842; <https://doi.org/10.1182/blood-2003-08-2894>. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=15059842](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15059842)
20. Hwang H-W, Park CY, Goodarzi H, Fak JJ, Mele A, Moore MJ, Saito Y, Darnell RB. PAPERCLIP identifies MicroRNA targets and a role of CstF64/64tau in promoting non-canonical poly(A) site usage. *Cell Rep* 2016; 15:423-35; PMID:27050522; <https://doi.org/10.1016/j.celrep.2016.03.023>. Available from <http://www.sciencedirect.com/science/article/pii/S2211124716302649>
  21. Kandala DT, Mohan N, A V, AP S, G R, Laishram RS. CstF-64 and 3'-UTR cis-element determine Star-PAP specificity for target mRNA selection by excluding PAP $\alpha$ . *Nucleic Acids Res* 2016; 44:811-23; PMID:26496945; <https://doi.org/10.1093/nar/gkv1074>. Available from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4737136/>
  22. Li X-Q, Du D. RNA polyadenylation sites on the genomes of microorganisms, animals, and plants. *PLoS One* 2013; 8:e79511; PMID:24260238; <https://doi.org/10.1371/journal.pone.0079511>. Available from <https://doi.org/10.1371%2Fjournal.pone.0079511>
  23. Gehring NH, Frede U, Neu-Yilik G, Hundsdorfer P, Vetter B, Hentze MW, Kulozik AE. Increased efficiency of mRNA 3[prime] end formation: A new genetic mechanism contributing to hereditary thrombophilia. *Nat Genet* 2001; 28:389-92; PMID:11443298; <https://doi.org/10.1038/ng578>
  24. Dalziel M, Nunes NM, Furger A. Two G-rich regulatory elements located adjacent to and 440 nucleotides downstream of the core poly(A) site of the intronless melanocortin receptor 1 gene are critical for efficient 3' end processing. *Mol Cell Biol* 2007; 27:1568-80; PMID:17189425; <https://doi.org/10.1128/MCB.01821-06>
  25. Arhin GK, Boots M, Bagga PS, Milcarek C, Wilusz J. Downstream sequence elements with different affinities for the hnRNP H/H' protein influence the processing efficiency of mammalian polyadenylation signals. *Nucleic Acids Res* 2002; 30:1842-50; PMID:11937639. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=11937639](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11937639)
  26. Dalziel M, Kolesnichenko M, das Neves RP, Iborra F, Goding C, Furger A.  $\alpha$ -MSH regulates intergenic splicing of MC1R and TUBB3 in human melanocytes. *Nucleic Acids Res* 2011; 39:2378-92; PMID:21071418; <https://doi.org/10.1093/nar/gkq1125>. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=21071418](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=21071418)
  27. Oberg D, Fay J, Lambkin H, Schwartz S. A downstream polyadenylation element in human papillomavirus type 16 L2 encodes multiple GGG motifs and interacts with hnRNP H. *J Virol* 2005; 79:9254-69; PMID:15994820; <https://doi.org/10.1128/JVI.79.14.9254-9269.2005>. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=15994820](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15994820)
  28. Oszolak F, Kapranov P, Foissac S, Kim SW, Fishilevich E, Monaghan AP, John B, Milos PM. Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell* 2011; 143:1018-29. Available from <http://linkinghub.elsevier.com/retrieve/pii/S0092867410013000>
  29. 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; 23:616-26; PMID:14749727; <https://doi.org/10.1038/sj.emboj.7600070>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/14749727>
  30. Lackford B, Yao C, Charles GM, Weng L, Zheng X, Choi EA, Xie X, Wan J, Xing Y, Freudenberg JM, et al. Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. *Embo J* 2014; 33:878-89; PMID:24596251; <https://doi.org/10.1002/embj.201386537>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/24596251>
  31. Li W, You B, Hoque M, Zheng D, Luo W, Ji Z, Park JY, Gunderson SI, Kalsotra A, Manley JL, 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:e1005166; PMID:25906188; <https://doi.org/10.1371/journal.pgen.1005166>. Available from <https://doi.org/10.1371%2Fjournal.pgen.1005166>
  32. Schek N, Cooke C, Alwine JC. Definition of the upstream efficiency element of the simian virus 40 late polyadenylation signal by using in vitro analyses. *Mol Cell Biol* 1992; 12:5386-93; PMID:1333042. Available from <http://mcb.asm.org/content/12/12/5386.abstract>
  33. Rothnie HM, Reid J, Hohn T. The contribution of AAUAAA and the upstream element UUUGUA to the efficiency of mRNA 3'-end formation in plants. *EMBO J* 1994; 13:2200-10; PMID:8187773. Available from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC395075/>
  34. Valsamakis A, Schek N, Alwine JC. Elements upstream of the AAUAAA within the human immunodeficiency virus polyadenylation signal are required for efficient polyadenylation in vitro. *Mol Cell Biol* 1992; 12:3699-705; PMID:1508176. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=1508176](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1508176)
  35. Gilmartin GM, Fleming ES, Oetjen J, Graveley BR. CPSF recognition of an HIV-1 mRNA 3'-processing enhancer: Multiple sequence contacts involved in poly(A) site definition. *Genes Dev* 1995; 9:72-83; PMID:7828853. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=7828853](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7828853)
  36. Brackenridge S, Proudfoot NJ. Recruitment of a basal polyadenylation factor by the upstream sequence element of the human lamin B2 polyadenylation signal. *Mol Cell Biol* 2000; 20:2660-9; PMID:10733568. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=10733568](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10733568)
  37. Moreira A, Takagaki Y, Brackenridge S, Wollerton M, Manley JL, Proudfoot NJ. The upstream sequence element of the C2 complement poly(A) signal activates mRNA 3' end formation by two distinct mechanisms. *Genes Dev* 1998; 12:2522-34; PMID:9716405. Available from <http://www.ncbi.nlm.nih.gov/pubmed/9716405>
  38. Danckwardt S, Kaufmann I, Gentzel M, Foerster KU, Gantzer AS, Gehring NH, Neu-Yilik G, Bork P, Keller W, Wilm M, et al. Splicing factors stimulate polyadenylation via USEs at non-canonical 3' end formation signals. *Embo J* 2007; 26:2658-69; PMID:17464285; <https://doi.org/10.1038/sj.emboj.7601699>. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=17464285](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17464285)
  39. Natalizio BJ, Muniz LC, Arhin GK, Wilusz J, Lutz CS. Upstream elements present in the 3'-untranslated region of collagen genes influence the processing efficiency of overlapping polyadenylation signals. *J Biol Chem* 2002; 277:42733-40; PMID:12200454; <https://doi.org/10.1074/jbc.M208070200>. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=12200454](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12200454)
  40. Hall-Pogar T, Zhang H, Tian B, Lutz CS. Alternative polyadenylation of cyclooxygenase-2. *Nucleic Acids Res* 2005; 33:2565-79. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=15872218](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15872218)
  41. Hall-Pogar T, Liang S, Hague LK, Lutz CS. Specific trans-acting proteins interact with auxiliary RNA polyadenylation elements in the COX-2 3'-UTR. *RNA* 2007; 13:1103-15; PMID:17507659; <https://doi.org/10.1261/rna.577707>. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=17507659](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17507659)
  42. Brown KM, Gilmartin GM. A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. *Mol Cell* 2003; 12:1467-76; PMID:14690600. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=14690600](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14690600)
  43. 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; 19:1315-27; PMID:15937220; <https://doi.org/10.1101/gad.1298605>. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=15937220](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15937220)

44. Masamha CP, Xia Z, Yang J, Albrecht TR, Li M, Shyu AB, Li W, Wagner EJ. CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. *Nature* 2014; 510:412-6; PMID:24814343; <https://doi.org/10.1038/nature13261>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/24814343>
45. Gruber AR, Martin G, Keller W, Zavolan M. Cleavage factor Im is a key regulator of 3' UTR length. *RNA Biol* 2012; 9:1405-12; PMID:23187700; <https://doi.org/10.4161/rna.22570>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/23187700>
46. Shi Y. Alternative polyadenylation: New insights from global analyses. *RNA* 2012; 18:2105-17; PMID:23097429; <https://doi.org/10.1261/rna.035899.112>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/23097429>
47. Mandel CR, Bai Y, Tong L. Protein factors in pre-mRNA 3'-end processing. *Cell Mol Life Sci* 2008; 65:1099-122; PMID:18158581; <https://doi.org/10.1007/s00018-007-7474-3>
48. Gilmartin GM, Nevins JR. Molecular analyses of two poly(A) site-processing factors that determine the recognition and efficiency of cleavage of the pre-mRNA. *Mol Cell Biol* 1991; 11:2432-8; PMID:2017162. Available from <http://www.ncbi.nlm.nih.gov/pubmed/2017162>
49. Schonemann L, Kuhn U, Martin G, Schafer P, Gruber AR, Keller W, Zavolan M, Wahle E. Reconstitution of CPSF active in polyadenylation: Recognition of the polyadenylation signal by WDR33. *Genes Dev* 2014; 28:2381-93; PMID:25301781; <https://doi.org/10.1101/gad.250985.114>; Available from <http://www.ncbi.nlm.nih.gov/pubmed/25301781>
50. Sullivan KD, Steiniger M, Marzluff WF. A core complex of CPSF73, CPSF100, and symplekin may form two different cleavage factors for processing of Poly(A) and Histone mRNAs. *Mol Cell* 2009; 34:322-32; PMID:19450530; <https://doi.org/10.1016/j.molcel.2009.04.024>. Available from <http://www.sciencedirect.com/science/article/pii/S1097276509002779>
51. Ohnacker M, Barabino SM, Preker PJ, Keller W. The WD-repeat protein pfs2p bridges two essential factors within the yeast pre-mRNA 3'-end-processing complex. *Embo J* 2000; 19:37-47; PMID:10619842; <https://doi.org/10.1093/emboj/19.1.37>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/10619842>
52. Shi Y, Di Giammartino DC, Taylor D, Sarkeshik A, Rice WJ, Yates Iii JR, Frank J, Manley JL. Molecular architecture of the human Pre-mRNA 3' processing complex. *Mol Cell* 2009; 33:365-76; PMID:19217410; <https://doi.org/10.1016/j.molcel.2008.12.028>. Available from <http://www.sciencedirect.com/science/article/B6WSR-4VKNNJ6-9/2/8e2b5763dacf d270ed74a12513447662>
53. Chan SL, Huppertz I, Yao C, Weng L, Moresco JJ, Yates 3rd JR, Ule J, Manley JL, Shi Y. CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3' processing. *Genes Dev* 2014; 28:2370-80; PMID:25301780; <https://doi.org/10.1101/gad.250993.114>; Available from <http://www.ncbi.nlm.nih.gov/pubmed/25301780>
54. Lau C, Bachorik JL, Dreyfuss G. Gemin5-snRNA interaction reveals an RNA binding function for WD repeat domains. *Nat Struct Mol Biol* 2009; 16:486-91; PMID:19377484; <https://doi.org/10.1038/nsmb.1584>
55. Moore CL, Chen J, Whoriskey J. Two proteins crosslinked to RNA containing the adenovirus L3 poly(A) site require the AAUAAA sequence for binding. *EMBO J* 1988; 7:3159-69; PMID:3181133. Available from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC454706/>
56. Delaney KJ, Xu R, Zhang J, Li QQ, Yun KY, Falcone DL, Hunt AG. Calmodulin interacts with and regulates the RNA-binding activity of an Arabidopsis polyadenylation factor subunit. *Plant Physiol* 2006; 140:1507-21; PMID:16500995; <https://doi.org/10.1104/pp.105.070672>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/16500995>
57. Thomas PE, Wu X, Liu M, Gaffney B, Ji G, Li QQ, Hunt AG. Genome-Wide control of polyadenylation site choice by CPSF30 in Arabidopsis. *Plant Cell* 2012; 24:4376-88; PMID:23136375; <https://doi.org/10.1105/tpc.112.096107>; Available from <http://www.ncbi.nlm.nih.gov/pubmed/23136375>
58. Shimberg GD, Michalek JL, Oluyadi AA, Rodrigues A V, Zucconi BE, Neu HM, Ghosh S, Sureschandra K, Wilson GM, Stemmler TL, et al. Cleavage and polyadenylation specificity factor 30: An RNA-binding zinc-finger protein with an unexpected 2Fe-2S cluster. *Proc Natl Acad Sci* 2016; 113:4700-5; PMID:27071088; <https://doi.org/10.1073/pnas.1517620113>. Available from <http://www.pnas.org/content/113/17/4700.abstract>
59. Bienroth S, Keller W, Wahle E. Assembly of a processive messenger RNA polyadenylation complex. *Embo J* 1993; 12:585-94; PMID:8440247. Available from <http://www.ncbi.nlm.nih.gov/pubmed/8440247>
60. Twu KY, Noah DL, Rao P, Kuo R-L, Krug RM. The CPSF30 binding site on the NS1A protein of influenza A virus is a potential antiviral target. *J Virol* 2006; 80:3957-65; PMID:16571812; <https://doi.org/10.1128/JVI.80.8.3957-3965.2006>. Available from <http://jvi.asm.org/content/80/8/3957.abstract>
61. Keller W, Bienroth S, Lang KM, Christofori G. Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. *Embo J* 1991; 10:4241-9; PMID:1756731. Available from <http://www.ncbi.nlm.nih.gov/pubmed/1756731>
62. Dichtl B, Blank D, Sadowski M, Hubner W, Weiser S, Keller W. Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination. *Embo J* 2002; 21:4125-35; PMID:12145212. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=12145212](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12145212)
63. Murthy KG, Manley JL. The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes Dev* 1995; 9:2672-83; PMID:7590244. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=7590244](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7590244)
64. Dominski Z, Yang X, Purdy M, Wagner EJ, Marzluff WF. A CPSF-73 homologue is required for cell cycle progression but not cell growth and interacts with a protein having features of CPSF-100. *Mol Cell Biol* 2005; 25:1489-500; PMID:15684398; <https://doi.org/10.1128/MCB.25.4.1489-1500.2005>. Available from <http://mcb.asm.org/content/25/4/1489.abstract>
65. Tran DDH, Saran S, Williamson AJK, Pierce A, Dittrich-Breiholz O, Wiehlmann L, Koch A, Whetton AD, Tamura T. THOC5 controls 3'-end-processing of immediate early genes via interaction with polyadenylation specific factor 100 (CPSF100). *Nucleic Acids Res* 2014; 42:12249-60; PMID:25274738; <https://doi.org/10.1093/nar/gku911>. Available from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4231767/>
66. Mandel CR, Kaneko S, Zhang H, Gebauer D, Vethantham V, Manley JL, Tong L. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* 2006; 444:953-6; PMID:17128255; <https://doi.org/10.1038/nature05363>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/17128255>
67. Kolev NG, Yario TA, Benson E, Steitz JA. Conserved motifs in both CPSF73 and CPSF100 are required to assemble the active endonuclease for histone mRNA 3'-end maturation. *EMBO Rep* 2008; 9:1013-8; PMID:18688255; <https://doi.org/10.1038/embor.2008.146>
68. Takagaki Y, Manley JL, MacDonald CC, Wilusz J, Shenk T. A multi-subunit factor, CstF, is required for polyadenylation of mammalian pre-mRNAs. *Genes Dev* 1990; 4:2112-20; PMID:1980119. Available from <http://www.ncbi.nlm.nih.gov/pubmed/1980119>
69. Takagaki Y, Ryner LC, Manley JL. Four factors are required for 3'-end cleavage of pre-mRNAs. *Genes Dev* 1989; 3:1711-24; PMID:2558045. Available from <http://www.ncbi.nlm.nih.gov/pubmed/2558045>
70. Gilmartin GM, Nevins JR. An ordered pathway of assembly of components required for polyadenylation site recognition and processing. *Genes Dev* 1989; 3:2180-90; PMID:2628166. Available from <http://www.ncbi.nlm.nih.gov/pubmed/2628166>
71. Takagaki Y, Manley JL. RNA recognition by the human polyadenylation factor CstF. *Mol Cell Biol* 1997; 17:3907-14; PMID:9199325. Available from <http://www.ncbi.nlm.nih.gov/pubmed/9199325>
72. Perez Canadillas JM, Varani G. Recognition of GU-rich polyadenylation regulatory elements by human CstF-64 protein. *Embo J* 2003; 22:2821-30; PMID:12773396; <https://doi.org/10.1093/>

- emboj/cdg259. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=12773396](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12773396)
73. Yao C, Biesinger J, Wan J, Weng L, Xing Y, Xie X, Shi Y. Transcriptome-wide analyses of CstF64–RNA interactions in global regulation of mRNA alternative polyadenylation. *Proc Natl Acad Sci* 2012; 109:18773-8; PMID:23112178; <https://doi.org/10.1073/pnas.1211101109>. Available from <http://www.pnas.org/content/109/46/18773.abstract>
  74. Takagaki Y, Manley JL. A polyadenylation factor subunit is the human homologue of the *Drosophila* suppressor of forked protein. *Nature* 1994; 372:471-4; PMID:7984242; <https://doi.org/10.1038/372471a0>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/7984242>
  75. 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; 25:863-75; PMID:17386263; <https://doi.org/10.1016/j.molcel.2007.01.034>. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=17386263](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17386263)
  76. Moreno-Morcillo M, Minvielle-SÁ@bastia L, Mackereth C, Fribourg S, Minvielle-Sébastien L, Mackereth C, Fribourg S. Hexameric architecture of CstF supported by CstF-50 homodimerization domain structure. *RNA* 2011; 17:412-8; PMID:21233223; <https://doi.org/10.1261/rna.2481011>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/21233223>
  77. McLauchlan J, Gaffney D, Whitton JL, Clements JB. The consensus sequence YGTGTTY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini. *Nucleic Acids Res* 1985; 13:1347-68; PMID:2987822. Available from <http://www.ncbi.nlm.nih.gov/pubmed/2987822>
  78. Kleiman FE, Manley JL. The BARD1–CstF-50 interaction links mRNA 3' end formation to DNA damage and tumor suppression. *Cell* 2001; 104:743-53; PMID:11257228. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=11257228](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11257228)
  79. Fong N, Bentley DL. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: Different functions for different segments of the CTD. *Genes Dev* 2001; 15:1783-95; PMID:11459828; <https://doi.org/10.1101/gad.889101>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/11459828>
  80. Yang Q, Coseno M, Gilmartin GM, Doublet S. Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping. *Structure* 2011; 19:368-77; PMID:21295486; <https://doi.org/10.1016/j.str.2010.12.021>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/21295486>
  81. Yang Q, Gilmartin GM, Doublet S. The structure of human cleavage factor I(m) hints at functions beyond UGUA-specific RNA binding: A role in alternative polyadenylation and a potential link to 5' capping and splicing. *RNA Biol* 2011; 8:748-53; PMID:21881408; <https://doi.org/10.4161/rna.8.5.16040>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/21881408>
  82. Rueggsegger 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; 271:6107-13; PMID:8626397. Available from <http://www.ncbi.nlm.nih.gov/pubmed/8626397>
  83. Kubo T, Wada T, Yamaguchi Y, Shimizu A, Handa H. Knock-down of 25 kDa subunit of cleavage factor Im in HeLa cells alters alternative polyadenylation within 3'-UTRs. *Nucleic Acids Res* 2006; 34:6264-71; PMID:17098938 PMID: PMC1669743 <https://doi.org/10.1093/nar/gkl794>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/17098938>
  84. Sartini BL, Wang H, Wang W, Millette CF, Kilpatrick DL. Pre-Messenger RNA Cleavage Factor I (CFIm): Potential role in alternative polyadenylation during spermatogenesis. *Biol Reprod* 2008; 78:472-82; PMID:18032416; <https://doi.org/10.1095/biolreprod.107.064774>. Available from <http://www.biolreprod.org/content/78/3/472.abstract>
  85. Katahira J, Okuzaki D, Inoue H, Yoneda Y, Maehara K, Ohkawa Y. Human TREX component Thoc5 affects alternative polyadenylation site choice by recruiting mammalian cleavage factor I. *Nucleic Acids Res* 2013; 41:7060-72; PMID:23685434; <https://doi.org/10.1093/nar/gkt414>. Available from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3737531&tool=pmcentrez&rendertype=abstract>
  86. de Vries H, Rueggsegger U, Hubner W, Friedlein A, Langen H, Keller W. Human pre-mRNA cleavage factor II(m) contains homologs of yeast proteins and bridges two other cleavage factors. *Embo J* 2000; 19:5895-904; PMID:11060040; <https://doi.org/10.1093/emboj/19.21.5895>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/11060040>
  87. West S, Proudfoot NJ. Human Pcf11 enhances degradation of RNA polymerase II-associated nascent RNA and transcriptional termination. *Nucleic Acids Res* 2008; 36:905-14; PMID:18086705; <https://doi.org/10.1093/nar/gkm1112>. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=18086705](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18086705)
  88. Raabe T, Bollum FJ, Manley JL. Primary structure and expression of bovine poly(A) polymerase. *Nature* 1991; 353:229-34; PMID:1896071; <https://doi.org/10.1038/353229a0>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/1896071>
  89. Wahle E, Martin G, Schiltz E, Keller W. Isolation and expression of cDNA clones encoding mammalian poly(A) polymerase. *Embo J* 1991; 10:4251-7; PMID:1756732. Available from <http://www.ncbi.nlm.nih.gov/pubmed/1756732>
  90. Wahle E. Purification and characterization of a mammalian polyadenylate polymerase involved in the 3' end processing of messenger RNA precursors. *J Biol Chem* 1991; 266:3131-9; PMID:1993684. Available from <http://www.ncbi.nlm.nih.gov/pubmed/1993684>
  91. Laishram RS. Poly(A) polymerase (PAP) diversity in gene expression–Star-PAP vs canonical PAP. *FEBS Lett* 2014; 588:2185-97; PMID:24873880; <https://doi.org/10.1016/j.febslet.2014.05.029>
  92. Deo RC, Bonanno JB, Sonenberg N, Burley SK. Recognition of polyadenylated RNA by the poly(A)-binding protein. *Cell* 1999; 98:835-45; PMID:10499800. Available from <http://www.ncbi.nlm.nih.gov/pubmed/10499800>
  93. Wahle E. A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell* 1991; 66:759-68; PMID:1878970. Available from <http://www.ncbi.nlm.nih.gov/pubmed/1878970>
  94. Kuhn U, Gundel M, Knoth A, Kerwitz Y, Rudel S, Wahle E, Kühn U, Gündel M, Knoth A, Kerwitz Y, et al. 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; 284:22803-14; PMID:19509282; <https://doi.org/10.1074/jbc.M109.018226>. Available from <http://www.jbc.org/content/284/34/22803.abstract>
  95. Chang H, Lim J, Ha M, Kim VN. TAIL-seq: Genome-wide determination of Poly(A) tail length and 3' end modifications. *Mol Cell* 2014; 53:1044-52; PMID:24582499; <https://doi.org/10.1016/j.molcel.2014.02.007>. Available from <http://www.sciencedirect.com/science/article/pii/S109727651400121X>
  96. Jenal M, Elkon R, Loayza-Puch F, van Haaften G, Kühn U, Menzies FM, Oude Vrielink JAF, Bos AJ, Drost J, Rooijers K, et al. The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* 2012; 149:538-53; PMID:22502866; <https://doi.org/10.1016/j.cell.2012.03.022>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/22502866>
  97. de Klerk E, Venema A, Anvar SY, Goeman JJ, Hu O, Trollet C, Dickson G, den Dunnen JT, van der Maarel SM, Raz V, et al. Poly(A) binding protein nuclear 1 levels affect alternative polyadenylation. *Nucleic Acids Res* 2012; 40:9089-101; PMID:22772983; <https://doi.org/10.1093/nar/gks655>. Available from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3467053&tool=pmcentrez&rendertype=abstract>
  98. Takagaki Y, Manley JL. Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol*

- Cell Biol 2000; 20:1515-25; PMID:10669729. Available from <http://www.ncbi.nlm.nih.gov/pubmed/10669729>
99. Xing H, Mayhew CN, Cullen KE, Park-Sarge OK, Sarge KD. HSF1 modulation of Hsp70 mRNA polyadenylation via interaction with symplekin. *J Biol Chem* 2004; 279:10551-5; PMID:14707147; <https://doi.org/10.1074/jbc.M311719200>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/14707147>
  100. Hirose Y, Manley JL. RNA polymerase II is an essential mRNA polyadenylation factor. *Nature* 1998; 395:93-6. Available from <http://www.ncbi.nlm.nih.gov/pubmed/9738505>
  101. Zaborowska J, Egloff S, Murphy S. The pol II CTD: New twists in the tail. *Nat Struct Mol Biol* 2016; 23:771-7; PMID:27605205; <https://doi.org/10.1038/nsmb.3285>
  102. Chao LC, Jamil A, Kim SJ, Huang L, Martinson HG. Assembly of the cleavage and polyadenylation apparatus requires about 10 seconds in vivo and is faster for strong than for weak poly(A) sites. *Mol Cell Biol* 1999; 19:5588-600; PMID:10409748. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=10409748](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10409748)
  103. Hart RP, McDevitt MA, Nevins JR. Poly(A) site cleavage in a HeLa nuclear extract is dependent on downstream sequences. *Cell* 1985; 43:677-83; PMID:2866847. Available from <http://www.ncbi.nlm.nih.gov/pubmed/2866847>
  104. Flaherty SM, Fortes P, Izaurrealde E, Mattaj IW, Gilmartin GM. Participation of the nuclear cap binding complex in pre-mRNA 3' processing. *Proc Natl Acad Sci* 1997; 94:11893-8; PMID:9342333. Available from <http://www.pnas.org/content/94/22/11893.abstract>
  105. Pandey NB, Chodchoy N, Liu TJ, Marzluff WF. Introns in histone genes alter the distribution of 3' ends. *Nucleic Acids Res* 1990; 18:3161-70; PMID:2356116. Available from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC330919/>
  106. Romeo V, Schümperli D. Cycling in the nucleus: Regulation of RNA 3' processing and nuclear organization of replication-dependent histone genes. *Curr Opin Cell Biol* 2016; 40:23-31; PMID:26895140; <https://doi.org/10.1016/j.ccb.2016.01.015>
  107. Niwa M, Rose SD, Berget SM. In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes Dev* 1990; 4:1552-9; PMID:1701407. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=1701407](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1701407)
  108. Niwa M, Berget SM. Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes Dev* 1991; 5:2086-95; PMID:1657710. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=1657710](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1657710)
  109. Dye MJ, Proudfoot NJ. Terminal exon definition occurs cotranscriptionally and promotes termination of RNA polymerase II. *Mol Cell* 1999; 3:371-8; PMID:10198639. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=10198639](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10198639)
  110. Davidson L, West S. Splicing-coupled 3' end formation requires a terminal splice acceptor site, but not intron excision. *Nucleic Acids Res* 2013; 41:7101-14; PMID:23716637; <https://doi.org/10.1093/nar/gkt446>. Available from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3737548/>
  111. Millevoi S, Loulergue C, Dettwiler S, Karaa SZ, Keller W, Antoniou M, Vagner S. An interaction between U2AF 65 and CF I(m) links the splicing and 3' end processing machineries. *EMBO J* 2006; 25:4854-64; PMID:17024186; <https://doi.org/10.1038/sj.emboj.7601331>
  112. Millevoi S, Geraghty F, Idowu B, Tam JL, Antoniou M, Vagner S. A novel function for the U2AF 65 splicing factor in promoting pre-mRNA 3'-end processing. *EMBO Rep* 2002; 3:869-74; PMID:12189174; <https://doi.org/10.1093/embo-reports/kvf173>. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=12189174](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12189174)
  113. Vagner S, Vagner C, Mattaj IW. The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing. *Genes Dev* 2000; 14:403-13; PMID:10691733. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=10691733](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10691733)
  114. Kyburz A, Friedlein A, Langen H, Keller W. Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of Pre-mRNA 3' end processing and splicing. *Mol Cell* 2006; 23:195-205; PMID:16857586; <https://doi.org/10.1016/j.molcel.2006.05.037>
  115. Misra A, Ou J, Zhu LJ, Green MR. Global promotion of alternative internal exon usage by mRNA 3' end formation factors. *Mol Cell* 2014; 58:819-31. Available from <http://www.sciencedirect.com/science/article/pii/S1097276515001811>
  116. Furth PA, Choe WT, Rex JH, Byrne JC, Baker CC. Sequences homologous to 5' splice sites are required for the inhibitory activity of papillomavirus late 3' untranslated regions. *Mol Cell Biol* 1994; 14:5278-89; PMID:8035806. Available from <http://www.ncbi.nlm.nih.gov/pubmed/8035806>
  117. Johansson C, Schwartz S. Regulation of human papillomavirus gene expression by splicing and polyadenylation. *Nat Rev Microbiol* 2013; 11:239-51; PMID:23474685; <https://doi.org/10.1038/nrmicro2984>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/23474685>
  118. Vagner S, Rueggsegger U, Gunderson SI, Keller W, Mattaj IW. Position-dependent inhibition of the cleavage step of pre-mRNA 3'-end processing by U1 snRNP. *RNA* 2000; 6:178-88; PMID:10688357. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=10688357](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10688357)
  119. Ashe MP, Furger A, Proudfoot NJ. Stem-loop 1 of the U1 snRNP plays a critical role in the suppression of HIV-1 polyadenylation. *RNA* 2000; 6:170-7; PMID:10688356
  120. Ashe MP, Griffin P, James W, Proudfoot NJ. Poly(A) site selection in the HIV-1 provirus: Inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. *Genes Dev* 1995; 9:3008-25; PMID:7498796. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=7498796](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7498796)
  121. Haenni S, Sharpe HE, Gravato noble M, Zechner K, Browne C, Hodgkin J, Furger A. Regulation of transcription termination in the nematode *Caenorhabditis elegans*. *Nucleic Acids Res* 2009; 37:6723-36; PMID:19740764; <https://doi.org/10.1093/nar/gkp744>
  122. Kaida D, Berg MG, Younis I, Kasim M, Singh LN, Wan L, Dreyfuss G. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* 2010; 468:664-8; PMID:20881964; <https://doi.org/10.1038/nature09479>. Available from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2996489&tool=pmcentrez&rendertype=abstract>
  123. Elkon R, Ugalde AP, Agami R. Alternative cleavage and polyadenylation: Extent, regulation and function. *Nat Rev Genet* 2013; 14:496-506; PMID:23774734; <https://doi.org/10.1038/nrg3482>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/23774734>
  124. Devany E, Park JY, Murphy MR, Zakusilo G, Baquero J, Zhang X, Hoque M, Tian B, Kleiman FE. Intronic cleavage and polyadenylation regulates gene expression during DNA damage response through U1 snRNA. *Cell Discov* 2016; 2:16013; PMID:27462460; <https://doi.org/10.1038/celldisc.2016.13>. Available from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4906801/>
  125. Kerwitz Y, Kuhn U, Lilie H, Knoth A, Scheuermann T, Friedrich H, Schwarz E, Wahle E. Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *Embo J* 2003; 22:3705-14; PMID:12853485; <https://doi.org/10.1093/emboj/cdg347>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/12853485>
  126. Brown CE, Sachs AB. Poly(A) tail length control in *Saccharomyces cerevisiae* occurs by message-specific deadenylation. *Mol Cell Biol* 1998; 18:6548-59; PMID:9774670. Available from <http://mcb.asm.org/content/18/11/6548.abstract>
  127. Chen Z, Li Y, Krug RM. Influenza A virus NS1 protein targets poly(A)-binding protein II of the cellular 3'-end processing machinery. *Embo J* 1999; 18:2273-83; PMID:10205180; <https://doi.org/10.1093/emboj/18.8.2273>
  128. Apponi LH, Leung SW, Williams KR, Valentini SR, Corbett AH, Pavlath GK. Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis. *Hum Mol Genet* 2010;

- 19:1058-65; PMID:20035013; <https://doi.org/10.1093/hmg/ddp569>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/20035013>
129. Saguez C, Schmid M, Olesen JR, Ghazy MA, Qu X, Poulsen MB, Nasser T, Moore C, Jensen TH. Nuclear mRNA surveillance in THO/sub2 mutants is triggered by inefficient polyadenylation. *Mol Cell* 2008; 31:91-103; PMID:18614048; <https://doi.org/10.1016/j.molcel.2008.04.030>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/18614048>
  130. Lee YJ, Glaunsinger BA. Aberrant herpesvirus-induced polyadenylation correlates with cellular messenger RNA destruction. *PLoS Biol* 2009; 7:e1000107; PMID:19468299; <https://doi.org/10.1371/journal.pbio.1000107>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/19468299>
  131. Jensen TH, Patricio K, McCarthy T, Rosbash M. A block to mRNA nuclear export in *S. cerevisiae* leads to hyperadenylation of transcripts that accumulate at the site of transcription. *Mol Cell* 2001; 7:887-98; PMID:11336711
  132. Sagawa F, Ibrahim H, Morrison AL, Wilusz CJ, Wilusz J. Nucleophosmin deposition during mRNA 3' end processing influences poly(A) tail length. *Embo J* 2011; 30:3994-4005; PMID:21822216; <https://doi.org/10.1038/emboj.2011.272>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/21822216>
  133. Jalkanen AL, Coleman SJ, Wilusz J. Determinants and implications of mRNA poly(A) tail size—Does this protein make my tail look big? *Semin Cell Dev Biol* 2014; 34:24-32; PMID:24910447; <https://doi.org/10.1016/j.semcdb.2014.05.018>
  134. Derti A, Garrett-Engele P, Macisaac KD, Stevens RC, Sriram S, Chen R, Rohl Ca, Johnson JM, Babak T. A quantitative atlas of polyadenylation in five mammals. *Genome Res* 2012; 22:1173-83; PMID:22454233; <https://doi.org/10.1101/gr.132563.111>
  135. You L, Wu J, Feng Y, Fu Y, Guo Y, Long L, Zhang H, Luan Y, Tian P, Chen L, et al. APASdb: A database describing alternative poly(A) sites and selection of heterogeneous cleavage sites downstream of poly(A) signals. *Nucleic Acids Res* 2015; 43:D59-67; PMID:25378337; <https://doi.org/10.1093/nar/gku1076>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/25378337>
  136. Jan CH, Friedman RC, Ruby JG, Bartel DP. Formation, regulation and evolution of *Caenorhabditis elegans* 3'UTRs. *Nature* 2011; 469:97-101; PMID:21085120; <https://doi.org/10.1038/nature09616>
  137. Hoque M, Ji Z, Zheng D, Luo W, Li W, You B, Park JY, Yehia G, Tian B. Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. *Nat Methods* 2013; 10:133-9; PMID:23241633; <https://doi.org/10.1038/nmeth.2288>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/23241633>
  138. Lin Y, Li Z, Ozsolak F. An in-depth map of polyadenylation sites in cancer. *Nucleic Acids Res* 2012; 40:8460-71; PMID:22753024; <https://doi.org/10.1093/nar/gks637>
  139. Beck AH, Weng Z, Witten DM, Zhu S, Foley JW, Lacroute P, Smith CL, Tibshirani R, van de Rijn M, Sidow A, et al. 3'-end Sequencing for expression quantification (3SEQ) from archival tumor samples. *PLoS One* 2010; 5:e8768; PMID:20098735; <https://doi.org/10.1371/journal.pone.0008768>
  140. Wang L, Dowell RD, Yi R. Genome-wide maps of polyadenylation reveal dynamic mRNA 3'-end formation in mammalian cell lineages. *RNA* 2013; 19:413-25; PMID:23325109; <https://doi.org/10.1261/rna.035360.112>. Available from <http://rnajournal.cshlp.org/content/19/3/413.abstract>
  141. Yoon OK, Hsu TY, Im JH, Brem RB. Genetics and regulatory impact of alternative polyadenylation in human B-lymphoblastoid cells. *PLoS Genet* 2012; 8:e1002882; PMID:22916029; <https://doi.org/10.1371/journal.pgen.1002882>
  142. Yoon OK, Brem RB. Noncanonical transcript forms in yeast and their regulation during environmental stress. *RNA* 2010; 16:1256-67; PMID:20421314; <https://doi.org/10.1261/rna.2038810>
  143. Haenni S, Ji Z, Hoque M, Rust N, Sharpe H, Eberhard R, Browne C, Hengartner MO, Mellor J, Tian B, et al. Analysis of *C. elegans* intestinal gene expression and polyadenylation by fluorescence-activated nuclei sorting and 3'-end-seq. *Nucleic Acids Res* 2012; 40:6304-18; PMID:22467213; <https://doi.org/10.1093/nar/gks282>
  144. Morris AR, Bos A, Diosdado B, Rooijers K, Elkon R, Bolijn AS, Carvalho B, Meijer GA, Agami R. Alternative cleavage and polyadenylation during colorectal cancer development. *Clin Cancer Res* 2012; 18:5256-66; PMID:22874640; <https://doi.org/10.1158/1078-0432.CCR-12-0543>
  145. Elkon R, Drost J, van Haften G, Jenal M, Schrier M, Oude Vrielink JA, Agami R. E2F mediates enhanced alternative polyadenylation in proliferation. *Genome Biol* 2012; 13:R59; PMID:22747694; <https://doi.org/10.1186/gb-2012-13-7-r59>
  146. Wilkening S, Pelechano V, Järvelin AI, Tekkedil MM, Anders S, Benes V, Steinmetz LM. An efficient method for genome-wide polyadenylation site mapping and RNA quantification. *Nucleic Acids Res* 2013; 41:1-8
  147. 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; 1:753-63; PMID:22813749; <https://doi.org/10.1016/j.celrep.2012.05.003>. Available from <http://www.sciencedirect.com/science/article/pii/S2211124712001271>
  148. Fox-Walsh K, Davis-Turak J, Zhou Y, Li H, Fu X-D. A multiplex RNA-seq strategy to profile poly(A+) RNA: Application to analysis of transcription response and 3' end formation. *Genomics* 2011; 98:266-71; PMID:21515359; <https://doi.org/10.1016/j.ygeno.2011.04.003>
  149. Shepard PJ, Choi E-AA, Lu J, Flanagan LA, Hertel KJ, Shi Y. Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. *RNA* 2011; 17:761-72; PMID:21343387; <https://doi.org/10.1261/rna.2581711>. Available from <http://rnajournal.cshlp.org/content/17/4/761.abstract>
  150. Moll P, Ante M, Seitz A, Reda T. QuantSeq 3' mRNA sequencing for RNA quantification. *Nat Methods* 2014; 11:25
  151. Fu Y, Sun Y, Li Y, Li J, Rao X, Chen C, Xu A. Differential genome-wide profiling of tandem 3' UTRs among human breast cancer and normal cells by high-throughput sequencing. *Genome Res* 2011; 21:741-7; PMID:21474764; <https://doi.org/10.1101/gr.115295.110>. Available from <http://genome.cshlp.org/content/21/5/741.abstract>
  152. Tian P, Li J, Liu X, Li Y, Chen M, Ma Y, Zheng YQ, Fu Y, Zou H. Tandem alternative polyadenylation events of genes in non-eosinophilic nasal polyp tissue identified by high-throughput sequencing analysis. *Int J Mol Med* 2014; 33:1423-30; PMID:24715051; <https://doi.org/10.3892/ijmm.2014.1734>
  153. 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; 320:1643-7; PMID:18566288; <https://doi.org/10.1126/science.1155390>
  154. Flavell SW, Kim T-KK, Gray JM, Harmin DA, Hemberg M, Hong EJ, Markenscoff-Papadimitriou E, Bear DM, Greenberg ME. Genome-wide analysis of MEF2 transcriptional program reveals synaptic target genes and neuronal activity-dependent polyadenylation site selection. *Neuron* 2008; 60:1022-38; PMID:19109909; <https://doi.org/10.1016/j.neuron.2008.11.029>. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=19109909](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19109909)
  155. Vilborg A, Passarelli MC, Yario TA, Tycowski KT, Steitz JA. Widespread inducible transcription downstream of human genes. *Mol Cell* 2015; 59:449-61; PMID:26190259; <https://doi.org/10.1016/j.molcel.2015.06.016>. Available from <http://www.sciencedirect.com/science/article/pii/S1097276515004554>
  156. Ni TK, Kuperwasser C. Premature polyadenylation of MAGI3 produces a dominantly-acting oncogene in human breast cancer. *Elife* 2016; 5:e14730; PMID:27205883; <https://doi.org/10.7554/eLife.14730>. Available from <https://dx.doi.org/10.7554/eLife.14730>
  157. Hollerer I, Curk T, Haase B, Benes V, Hauer C, Neu-Yilik G, Bhuvanagiri M, Hentze MW, Kulozik AE. The differential expression of alternatively polyadenylated transcripts is a common stress-induced response mechanism that modulates mammalian mRNA expression in a quantitative and qualitative fashion. *RNA* 2016; 22:1441-53; PMID:27407180; <https://doi.org/10.1261/rna.055657.115>; Available from <http://rnajournal.cshlp.org/content/early/2016/07/12/rna.055657.115.abstract>

158. 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; 106:7028-33; PMID:19372383; <https://doi.org/10.1073/pnas.0900028106>
159. 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:e8419; PMID:20037631; <https://doi.org/10.1371/journal.pone.0008419>
160. Li W, Park JY, Zheng D, Hoque M, Yehia G, Tian B. Alternative cleavage and polyadenylation in spermatogenesis connects chromatin regulation with post-transcriptional control. *BMC Biol* 2016; 14:6; PMID:26801249; <https://doi.org/10.1186/s12915-016-0229-6>. Available from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4724118/>
161. Miura P, Shenker S, Andreu-Agullo C, Westholm JO, Lai EC. Widespread and extensive lengthening of 3' UTRs in the mammalian brain. *Genome Res* 2013; 23:812-25; PMID:23520388; <https://doi.org/10.1101/gr.146886.112>; gr.146886.112-. Available from <http://genome.cshlp.org/content/early/2013/03/20/gr.146886.112.abstract>
162. Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue transcriptomes. *Nature* 2008; 456:470-6; PMID:18978772; <https://doi.org/10.1038/nature07509>
163. Smibert P, Miura P, Westholm JO, Shenker S, May G, Duff MO, Zhang D, Eads BD, Carlson J, Brown JB, et al. Global patterns of tissue-specific alternative polyadenylation in drosophila. *Cell Rep* 2012; 1:277-89; PMID:22685694; <https://doi.org/10.1016/j.celrep.2012.01.001>. Available from <http://linkinghub.elsevier.com/retrieve/pii/S2211124712000216>
164. Lianoglou S, Garg V, Yang JL, Leslie CS, Mayr C. Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes Dev* 2013; 27:2380-96; PMID:24145798; <https://doi.org/10.1101/gad.229328.113>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/24145798>
165. Wang Y, Loomis PA, Zinkowski RP, Binder LI. A novel tau transcript in cultured human neuroblastoma cells expressing nuclear tau. *J Cell Biol* 1993; 121:257-67; PMID:8468346
166. Liu D, Brockman JM, Dass B, Hutchins LN, Singh P, McCarrey JR, MacDonald CC, Graber JH. Systematic variation in mRNA 3'-processing signals during mouse spermatogenesis. *Nucleic Acids Res* 2007; 35:234-46; PMID:17158511; <https://doi.org/10.1093/nar/gkl919>. Available from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1802579&tool=pmcentrez&rendertype=abstract>
167. MacDonald CC, McMahon KW. Tissue-specific mechanisms of alternative polyadenylation: Testis, brain, and beyond. *Wiley Interdiscip Rev RNA* 2010; 1:494-501; PMID:21956945; <https://doi.org/10.1002/wrna.29>
168. Dass B, McDaniel L, Schultz RA, Attaya E, MacDonald CC. The gene CSTF2T, encoding the human variant CstF-64 polyadenylation protein  $\tau$ CstF-64, lacks introns and may be associated with male sterility. *Genomics* 2002; 80:509-14; PMID:12408968
169. Ara T, Lopez F, Ritchie W, Benech P, Gautheret D. Conservation of alternative polyadenylation patterns in mammalian genes. *BMC Genomics* 2006; 7:189; PMID:16872498; <https://doi.org/10.1186/1471-2164-7-189>
170. Moqtaderi Z, Geisberg J V, Jin Y, Fan X, Struhl K. Species-specific factors mediate extensive heterogeneity of mRNA 3' ends in yeasts. *Proc Natl Acad Sci U S A* 2013; 110:11073-8; PMID:23776204; <https://doi.org/10.1073/pnas.1309384110>
171. Velten L, Anders S, Pekowska A, Järvelin AI, Huber W, Pelechano V, Steinmetz LM. Single-cell polyadenylation site mapping reveals 3' isoform choice variability. *Mol Syst Biol* 2015; 11:812; PMID:26040288. Available from <http://msb.embopress.org/msb/11/6/812.full.pdf>
172. Mayr C, Bartel DP. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* 2009; 138:673-84; PMID:19703394; <https://doi.org/10.1016/j.cell.2009.06.016>
173. Singh P, Alley TL, Wright SM, Kamdar S, Schott W, Wilpan RY, Mills KD, Graber JH. Global changes in processing of mRNA 3' untranslated regions characterize clinically distinct cancer subtypes. *Cancer Res* 2009; 69:9422-30; PMID:19934316; <https://doi.org/10.1158/0008-5472.CAN-09-2236>
174. Xia Z, Donehower LA, Cooper TA, Neilson JR, Wheeler DA, Wagner EJ, Li W. Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types. *Nat Commun* 2014; 5:5274; PMID:25409906; <https://doi.org/10.1038/ncomms6274>
175. Rehfeld A, Plass M, Døssing K, Knigge U, Kjær A, Krogh A, Friis-Hansen L. Alternative polyadenylation of tumor suppressor genes in small intestinal neuroendocrine tumors. *Front Endocrinol (Lausanne)* 2014; 5:46; PMID:24782827; <https://doi.org/10.3389/fendo.2014.00046>
176. Wang X, Li M, Yin Y, Li L, Tao Y, Chen D, Li J, Han H, Hou Z, Zhang B, et al. Profiling of alternative polyadenylation sites in luminal B breast cancer using the SAPAS method. *Int J Mol Med* 2015; 35:39-50; PMID:25333330; <https://doi.org/10.3892/ijmm.2014.1973>. Available from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4249744/>
177. Lembo A, Di Cunto F, Provero P. Shortening of 3'UTRs correlates with poor prognosis in breast and lung cancer. *PLoS One* 2012; 7:e31129; PMID:22347440; <https://doi.org/10.1371/journal.pone.0031129>
178. Wang L, Hu X, Wang P, Shao Z-M. The 3'UTR signature defines a highly metastatic subgroup of triple-negative breast cancer. *Oncotarget* 2016; 7:59834-44; PMID:27494850; <https://doi.org/10.18632/oncotarget.10975>. Available from <http://www.impactjournals.com/oncotarget/index.php?journal=oncotarget&page=article&op=view&path%255B%255D=10975>
179. Bennett CL, Brunkow ME, Ramsdell F, O'Brian KC, Zhu Q, Fuleihan RL, Shigeoka A O, Ochs HD, Chance PF. A rare polyadenylation signal mutation of the FOXP3 gene (AAUAAA $\rightarrow$ AAUGAA) leads to the IPEX syndrome. *Immunogenetics* 2001; 53:435-9; PMID:11685453; <https://doi.org/10.1007/s002510100358>
180. Whitelaw E, Proudfoot N. Alpha-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human alpha 2 globin gene. *Embo J* 1986; 5:2915-22; PMID:3024968. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=3024968](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3024968)
181. Yasuda M, Shabbeer J, Osawa M, Desnick RJ. Fabry disease: Novel alpha-galactosidase A 3'-terminal mutations result in multiple transcripts due to aberrant 3'-end formation. *Am J Hum Genet* 2003; 73:162-73; PMID:12796853; <https://doi.org/10.1086/376608>
182. Wiestner A, Tehrani M, Chiorazzi M, Wright G, Gibellini F, Nakayama K, Liu H, Rosenwald A, Muller-Hermelink HK, Ott G, et al. Point mutations and genomic deletions in CCND1 create stable truncated cyclin D1 mRNAs that are associated with increased proliferation rate and shorter survival. *Blood* 2007; 109:4599-606; PMID:17299095; <https://doi.org/10.1182/blood-2006-08-039859>. Available from <http://bloodjournal.hematologylibrary.org/content/109/11/4599.abstract>
183. Brais B, Bouchard JP, Xie YG, Rochefort DL, Chrétien N, Tomé FM, Lafrenière RG, Rommens JM, Uyama E, Nohira O, et al. Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat Genet* 1998; 18:164-7; PMID:9462747; <https://doi.org/10.1038/ng0298-164>
184. Batra R, Charizanis K, Manchanda M, Mohan A, Li M, Finn DJ, Goodwin M, Zhang C, Sobczak K, Thornton CA, et al. Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-Mediated disease. *Mol Cell* 2014; 56:311-22; PMID:25263597; <https://doi.org/10.1016/j.molcel.2014.08.027>. Available from <http://www.sciencedirect.com/science/article/pii/S1097276514006820>
185. Prudencio M, Belzil VV, Batra R, Ross CA, Gendron TF, Pregent LJ, Murray ME, Overstreet KK, Piazza-Johnston AE, Desaro P, et al. Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat Neurosci* 2015; 18:1175-82; PMID:26192745; <https://doi.org/10.1038/nn.4065>

186. Zhou L, Liu F, Wang X, Ouyang G. The roles of microRNAs in the regulation of tumor metastasis. *Cell & Biosci* 2015; 5:1-12. Available from <https://doi.org/10.1186/s13578-015-0028-8>
187. Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nat Rev Cancer* 2006; 6:857-66
188. Liaw H-HH, Lin C-CC, Juan H-FF, Huang H-CC. Differential MicroRNA regulation correlates with alternative polyadenylation pattern between breast cancer and normal cells. *PLoS One* 2013; 8:e56958; PMID:23437281; <https://doi.org/10.1371/journal.pone.0056958>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/23437281>
189. Taliaferro JMM, Vidaki M, Oliveira R, Olson S, Zhan L, Saxena T, Wang ETT, Graveley BRR, Gertler FBB, Swanson MSS, et al. Distal alternative last exons localize mRNAs to neural projections. *Mol Cell* 2016; 61:821-33; PMID:26907613; <https://doi.org/10.1016/j.molcel.2016.01.020>
190. Kim S, Yamamoto J, Chen Y, Aida M, Wada T, Handa H, Yamaguchi Y. Evidence that cleavage factor Im is a heterotrimeric protein complex controlling alternative polyadenylation. *Genes Cells* 2010; 15:1003-13; PMID:20695905; <https://doi.org/10.1111/j.1365-2443.2010.01436.x>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/20695905>
191. Gennarino VA, Alcott CE, Chen C-A, Chaudhury A, Gillentine MA, Rosenfeld JA, Parikh S, Wheless JW, Roeder ER, Horovitz DDG, et al. NUDT21-spanning CNVs lead to neuropsychiatric disease and altered MeCP2 abundance via alternative polyadenylation. *Elife* 2015; 4:e10782. Available from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4586391/>
192. Shell SA, Hesse C, Morris SM, Milcarek C. Elevated levels of the 64-kDa cleavage stimulatory factor (CstF-64) in lipopolysaccharide-stimulated macrophages influence gene expression and induce alternative poly(A) site selection. *J Biol Chem* 2005; 280:39950-61; PMID:16207706; <https://doi.org/10.1074/jbc.M508848200>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/16207706>
193. Takagaki Y, Seipelt RL, Peterson ML, Manley JL. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell* 1996; 87:941-52; PMID:8945520. Available from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=8945520](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8945520)
194. Luo W, Ji Z, Pan Z, You B, Hoque M, Li W, Gunderson SI, Tian B. The conserved intronic cleavage and polyadenylation site of CstF-77 gene imparts control of 3' end processing activity through feedback autoregulation and by U1 snRNP. *PLoS Genet* 2013; 9:e1003613; PMID:23874216; <https://doi.org/10.1371/journal.pgen.1003613>
195. Shankarling GS, Macdonald CC. Polyadenylation site-specific differences in the activity of the neuronal  $\beta$ CstF-64 protein in PC12 cells. *Gene* 2013; 64:6-13
196. Simonelig M. PABPN1 shuts down alternative poly(A) sites. *Cell Res* 2012; 22:1419-21; PMID:22641371; <https://doi.org/10.1038/cr.2012.86>
197. Beaulieu YB, Kleinman CL, Landry-Voyer A-M, Majewski J, Bachand F. Polyadenylation-dependent control of long noncoding RNA expression by the poly(A)-binding protein nuclear 1. *PLoS Genet* 2012; 8:e1003078; PMID:23166521; <https://doi.org/10.1371/journal.pgen.1003078>. Available from [http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3499365&tool=pmcentrez&render\\_type=abstract](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3499365&tool=pmcentrez&render_type=abstract)
198. Ji X, Wan J, Vishnu M, Xing Y, Liebhaber SA. alphaCP Poly(C) binding proteins act as global regulators of alternative polyadenylation. *Mol Cell Biol* 2013; 33:2560-73. Available from <http://www.ncbi.nlm.nih.gov/pubmed/23629627>
199. Liu Y, Hu W, Murakawa Y, Yin J, Wang G, Landthaler M, Yan J. Cold-induced RNA-binding proteins regulate circadian gene expression by controlling alternative polyadenylation. *Sci Rep* 2013; 3:2054; PMID:23792593; <https://doi.org/10.1038/srep02054>. Available from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3690385&tool=pmcentrez&rendertype=abstract>
200. Bava F-A, Eliscovich C, Ferreira PG, Miñana B, Ben-Dov C, Guigó R, Valcárcel J, Méndez R. CPEB1 coordinates alternative 3'-UTR formation with translational regulation. *Nature* 2013; 1-7. Available from <http://www.ncbi.nlm.nih.gov/pubmed/23434754>
201. Oktaba K, Zhang W, Lotz TS, Jun DJ, Lemke SB, Ng SP, Esposito E, Levine M, Hilgers V. ELAV links paused Pol II to alternative polyadenylation in the drosophila nervous system. *Mol Cell* 2014; 1-8
202. Slobodin B, Agami R. Transcription initiation determines its end. *Mol Cell* 2015; 57:205-6; PMID:25616066; <https://doi.org/10.1016/j.molcel.2015.01.006>
203. Masuda A, Takeda J, Okuno T, Okamoto T, Ohkawara B, Ito M, Ishigaki S, Sobue G, Ohno K. Position-specific binding of FUS to nascent RNA regulates mRNA length. *Genes Dev* 2015; 29:1045-57; PMID:25995189; <https://doi.org/10.1101/gad.255737.114>. Available from <http://genesdev.cshlp.org/content/29/10/1045.abstract>
204. Alkan SA, Martincic K, Milcarek C. The hnRNPs F and H2 bind to similar sequences to influence gene expression. *Biochem J* 2006; 393:361-71; PMID:16171461; <https://doi.org/10.1042/BJ20050538>
205. Katz Y, Wang ET, Airoidi EM, Burge CB. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods* 2010; 7:1009-15; PMID:21057496; <https://doi.org/10.1038/nmeth.1528>
206. Bagga PS, Arhin GK, Wilusz J. DSEF-1 is a member of the hnRNP H family of RNA-binding proteins and stimulates pre-mRNA cleavage and polyadenylation in vitro. *Nucleic Acids Res* 1998; 26:5343-50; PMID:9826757
207. Naganuma T, Nakagawa S, Tanigawa A, Sasaki YF, Goshima N, Hirose T. Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *TL - 31. EMBO J* 2012; 31 VN-r:4020-34; PMID:22960638; <https://doi.org/10.1038/emboj.2012.251>. Available from [/Users/yurikoharigaya/Documents/ReadCube Media/Robert\\_9\\_21\\_12.pdf%5Cnhttps://doi.org/10.1038/emboj.2012.251](/Users/yurikoharigaya/Documents/ReadCube Media/Robert_9_21_12.pdf%5Cnhttps://doi.org/10.1038/emboj.2012.251)
208. Hung L-H, Heiner M, Hui J, Schreiner S, Benes V, Bindereif A. Diverse roles of hnRNP L in mammalian mRNA processing: A combined microarray and RNAi analysis. *RNA* 2008; 14:284-96; PMID:18073345; <https://doi.org/10.1261/rna.725208>
209. Barnhart MD, Moon SL, Emch AW, Wilusz CJ, Wilusz J. Changes in cellular mRNA stability, splicing, and polyadenylation through HuR protein sequestration by a cytoplasmic RNA virus. *Cell Rep* 2013; 1-9
210. Dai W, Zhang G, Makeyev EV. RNA-binding protein HuR autoregulates its expression by promoting alternative polyadenylation site usage. *Nucleic Acids Res* 2012; 40:787-800; PMID:21948791; <https://doi.org/10.1093/nar/gkr783>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/21948791>
211. Berkovits BD, Mayr C. Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. *Nature* 2015; 522:363-7; PMID:25896326; <https://doi.org/10.1038/nature14321>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/25896326>
212. Huang Y, Li W, Yao X, Lin Q, Yin J, Liang Y, Heiner M, Tian B, Hui J, Wang G. Mediator complex regulates alternative mRNA processing via the MED23 subunit. *Mol Cell* 2012; 45:459-69; PMID:22264826; <https://doi.org/10.1016/j.molcel.2011.12.022>
213. Nimura K, Yamamoto M, Takeichi M, Saga K, Takaoka K, Kawamura N, Nitta H, Nagano H, Ishino S, Tanaka T, et al. Regulation of alternative polyadenylation by Nkx2-5 and Xrn2 during mouse heart development. *Elife* 2016; 5:e16030; PMID:27331609; <https://doi.org/10.7554/eLife.16030>. Available from <https://dx.doi.org/10.7554/eLife.16030>
214. Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 2008; 456:464-9; PMID:18978773; <https://doi.org/10.1038/nature07488>
215. Yang Y, Li W, Hoque M, Hou L, Shen S, Tian B, Dynlacht BD. PAF complex plays novel subunit-specific roles in alternative cleavage and polyadenylation. *PLoS Genet* 2016; 12:e1005794; PMID:26765774; <https://doi.org/10.1371/journal.pgen.1005794>. Available from <https://doi.org/10.1371%252Fjournal.pgen.1005794>
216. Castelo-Branco P, Furger A, Wollerton M, Smith C, Moreira A, Proudfoot N. Polypyrimidine tract binding protein modulates efficiency of polyadenylation. *Mol Cell Biol* 2004; 24:4174-83; PMID:15121839

217. Di Giammartino DC, Li W, Ogami K, Yashinskii JJ, Hoque M, Tian B, Manley JL. RBBP6 isoforms regulate the human polyadenylation machinery and modulate expression of mRNAs with AU-rich 3' UTRs. *Genes Dev* 2014; 28:2248-60; PMID:25319826; <https://doi.org/10.1101/gad.245787.114>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/25319826>
218. McCracken S, Lamberon M, Benjamin J, Blencowe BJ. SRm160 splicing coactivator promotes transcript 3'-End cleavage. *Mol Cell Biol* 2002; 22:148-60; PMID:11739730
219. Müller-McNicoll M, Botti V, de Jesus Domingues AM, Brandl H, Schwich OD, Steiner MC, Curk T, Poser I, Zarnack K, Neugebauer KM. SR proteins are NXF1 adaptors that link alternative RNA processing to mRNA export. *Genes Dev* 2016; 30:553-66; PMID:26944680; <https://doi.org/10.1101/gad.276477.115>. Available from <http://genesdev.cshlp.org/content/30/5/553.abstract>
220. Koyama A, Sugai A, Kato T, Ishihara T, Shiga A, Toyoshima Y, Koyama M, Konno T, Hirokawa S, Yokoseki A, et al. Increased cytoplasmic TARDBP mRNA in affected spinal motor neurons in ALS caused by abnormal autoregulation of TDP-43. *Nucleic Acids Res* 2016; 44:5820-36; PMID:27257061; <https://doi.org/10.1093/nar/gkw499>. Available from <http://nar.oxfordjournals.org/content/44/12/5820.abstract>
221. Berg MG, Singh LN, Younis I, Liu Q, Pinto AM, Kaida D, Zhang Z, Cho S, Sherrill-Mix S, Wan L, et al. U1 snRNP determines mRNA length and regulates isoform expression. *Cell* 2012; 150:53-64; PMID:22770214; <https://doi.org/10.1016/j.cell.2012.05.029>
222. Lou H, Neugebauer K. Regulation of Alternative Polyadenylation by U1 snRNPs and SRp20. *Mol Cell Biol* 1998; 18:4977
223. Phillips C, Pachikara N, Gunderson SI. U1A inhibits cleavage at the immunoglobulin M heavy-chain secretory poly(A) site by binding between the two downstream GU-rich regions. *Mol Cell Biol* 2004; 24:6162-71; PMID:15226420; <https://doi.org/10.1128/MCB.24.14.6162-6171.2004>
224. Gunderson SI, Beyer K, Martin G, Keller W, Boelens WC, Mattaj JW. The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. *Cell* 1994; 76:531-41; PMID:8313473. Available from <http://www.sciencedirect.com/science/article/pii/0092867494901163>
225. West S, Proudfoot NJ. Transcriptional termination enhances protein expression in human cells. *Mol Cell* 2009; 33:354-64; PMID:19217409; <https://doi.org/10.1016/j.molcel.2009.01.008>. Available from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2706331/>
226. Mapendano CK, Lykke-Andersen S, Kjems J, Bertrand E, Jensen TH. Crosstalk between mRNA 3' End processing and transcription initiation. *Mol Cell* 2010; 40:410-22; PMID:21070967; <https://doi.org/10.1016/j.molcel.2010.10.012>
227. Hsin J-P, Manley JL. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev* 2012; 26:2119-37; PMID:23028141; <https://doi.org/10.1101/gad.200303.112>. Available from <http://genesdev.cshlp.org/content/26/19/2119.abstract>
228. Fusby B, Kim S, Erickson B, Kim H, Peterson ML, Bentley DL. Coordination of RNA Polymerase II Pausing and 3' end processing factor recruitment with alternative polyadenylation. *Mol Cell Biol* 2015; 36:295-303; PMID:26527620; <https://doi.org/10.1128/MCB.00898-15>
229. Rigo F, Kazerounia A, Nag A, Martinson HG. The RNA tether from the poly(A) signal to the polymerase mediates coupling of transcription to cleavage and polyadenylation. *Mol Cell* 2005; 20:733-45; PMID:16337597; <https://doi.org/10.1016/j.molcel.2005.09.026>
230. Pinto PB, Henriques T, Freitas MO, Martins T, Domingues RG, Wyrzykowska PS, Coelho P, Carmo AM, Sunkel CE, Proudfoot NJ, et al. RNA polymerase II kinetics in polo polyadenylation signal selection. *EMBO J* 2011; 30:2431-44; PMID:21602789; <https://doi.org/10.1038/emboj.2011.156>
231. Huang H, Chen J, Liu H, Sun X. The nucleosome regulates the usage of polyadenylation sites in the human genome. *BMC Genomics* 2013; 14:912; PMID:24365105; <https://doi.org/10.1186/1471-2164-14-912>. Available from <http://www.biomedcentral.com/1471-2164/14/912>
232. Spies N, Nielsen CB, Padgett RA, Burge CB. Biased chromatin signatures around polyadenylation sites and exons. *Mol Cell* 2009; 36:245-54; PMID:19854133; <https://doi.org/10.1016/j.molcel.2009.10.008>
233. Nagaike T, Logan C, Hotta I, Rozenblatt-Rosen O, Meyerson M, Manley JL. Transcriptional activators enhance polyadenylation of mRNA precursors. *Mol Cell* 2011; 41:409-18; PMID:21329879; <https://doi.org/10.1016/j.molcel.2011.01.022>. Available from <http://www.sciencedirect.com/science/article/pii/S1097276511000499>
234. Wood AJ, Schulz R, Woodfine K, Koltowska K, Beechey C V, Peters J, Bourc'his D, Oakey RJ, Bourc'his D, Oakey RJ. Regulation of alternative polyadenylation by genomic imprinting. *Genes Dev* 2008; 22:1141-6; PMID:18451104; <https://doi.org/10.1101/gad.473408>. Available from <http://genesdev.cshlp.org/content/22/9/1141.abstract>
235. Cowley M, Wood AJ, Bohm S, Schulz R, Oakey RJ. Epigenetic control of alternative mRNA processing at the imprinted Herc3/Nap115 locus. *Nucleic Acids Res* 2012; 40:8917-26; PMID:22790983; <https://doi.org/10.1093/nar/gks654>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/22790983>
236. Peng S, Chen CYA, Xu N, Shyu AB. RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein. *EMBO J* 1998; 17:3461-70; PMID:9628881; <https://doi.org/10.1093/emboj/17.12.3461>
237. Park-Lee S, Kim S, Laird-Offringa IA. Characterization of the Interaction between Neuronal RNA-binding Protein HuD and AU-rich RNA. *J Biol Chem* 2003; 278:39801-8; PMID:12900401; <https://doi.org/10.1074/jbc.M307105200>. Available from <http://www.jbc.org/content/278/41/39801.abstract>
238. Abdelmohsen K, Hutchison ER, Lee EK, Kuwano Y, Kim MM, Masuda K, Srikantan S, Subaran SS, Marasa BS, Mattson MP, et al. miR-375 inhibits differentiation of neurites by lowering HuD levels. *Mol Cell Biol* 2010; 30:4197-210; PMID:20584986; <https://doi.org/10.1128/MCB.00316-10>. Available from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2937556&tool=pmcentrez&rendertype=abstract>
239. Aranda-Abreu GE, Behar L, Chung S, Furneaux H, Ginzburg I. Embryonic lethal abnormal Vision-Like RNA-Binding proteins regulate Neurite outgrowth and Tau expression in PC12 cells. *J Neurosci* 1999; 19:6907-17; PMID:10436048. Available from <http://www.jneurosci.org/content/19/16/6907.abstract>
240. Beisang D, Bohjanen PR. Perspectives on the ARE as it turns 25 years old. *Wiley Interdiscip Rev RNA* 2012; 3:719-31; PMID:22733578; <https://doi.org/10.1002/wrna.1125>
241. Bakheet T, Williams BRG, Khabar KS. a. ARED 3.0: The large and diverse AU-rich transcriptome. *Nucleic Acids Res* 2006; 34:D111-4
242. Yano M, Okano HJ, Okano H. Involvement of Hu and heterogeneous nuclear Ribonucleoprotein K in Neuronal differentiation through p21 mRNA Post-transcriptional regulation. *J Biol Chem* 2005; 280:12690-9; PMID:15671036; <https://doi.org/10.1074/jbc.M411119200>. Available from <http://www.jbc.org/content/280/13/12690.abstract>
243. Rattenbacher B, Beisang D, Wiesner DL, Jeschke JC, von Hohenberg M, St Louis-Vlasova IA, Bohjanen PR. Analysis of CUGBP1 targets identifies GU-repeat sequences that mediate rapid mRNA decay. *Mol Cell Biol* 2010; 30:3970-80; PMID:20547756; <https://doi.org/10.1128/MCB.00624-10>
244. Sawicka K, Bushell M, Spriggs KA, Willis AE. Polypyrimidine-tract-binding protein: A multifunctional RNA-binding protein. *Biochem Soc Trans* 2008; 36:641-7; PMID:18631133; <https://doi.org/10.1042/BST0360641>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/18631133>
245. Lee D-H, Lim M-H, Youn D-Y, Jung SE, Ahn YS, Tsujimoto Y, Lee J-H. hnRNP L binds to CA repeats in the 3'UTR of bcl-2 mRNA. *Biochem Biophys Res Commun* 2009; 382:583-7; PMID:19298794; <https://doi.org/10.1016/j.bbrc.2009.03.069>
246. Friedman RC, Farh KK-H, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009; 19:92-105; PMID:18955434; <https://doi.org/10.1101/gr.082701.108>

247. Guo H, Ingolia NT, Weissman JS, Bartel DP. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 2010; 466:835-40; PMID:20703300; <https://doi.org/10.1038/nature09267>
248. Legendre M, Ritchie W, Lopez F, Gautheret D. Differential repression of alternative transcripts: A screen for miRNA targets. *PLoS Comput Biol* 2006; 2:e43; PMID:16699595; <https://doi.org/10.1371/journal.pcbi.0020043>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/16699595>
249. Boutet SC, Cheung TH, Quach NL, Liu L, Prescott SL, Edalati A, Iori K, Rando TA. Alternative polyadenylation mediates microRNA regulation of muscle stem cell function. *Cell Stem Cell* 2012; 10:327-36; PMID:22385659; <https://doi.org/10.1016/j.stem.2012.01.017>. Available from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3306803&tool=pmcentrez&rendertype=abstract>
250. Kreth S, Limbeck E, Hinske LC, Schütz S V, Thon N, Hoefig K, Egensperger R, Kreth FW. In human glioblastomas transcript elongation by alternative polyadenylation and miRNA targeting is a potent mechanism of MGMT silencing. *Acta Neuropathol* 2013; 125:671-81; PMID:23340988; <https://doi.org/10.1007/s00401-013-1081-1>
251. Akman BH, Can T, Erson-Bensan AE. Estrogen-induced upregulation and 3'-UTR shortening of CDC6. *Nucleic Acids Res* 2012; 40:10679-88; PMID:22977174; <https://doi.org/10.1093/nar/gks855>
252. Hinske LC, Galante PAF, Limbeck E, Möhnle P, Parmigiani RB, Ohno-Machado L, Camargo AA, Kreth S. Alternative Polyadenylation allows differential negative feedback of human miRNA miR-579 on Its Host Gene ZFR. *PLoS One* 2015; 10:e0121507; PMID:25799583; <https://doi.org/10.1371/journal.pone.0121507>
253. Lazarov ME, Martin MM, Willardson BM, Elton TS. Human phosphatase-like protein (hPhLP) messenger RNA stability is regulated by cis-acting instability elements present in the 3'-untranslated region. *Biochim Biophys Acta* 1999; 1446:253-64; PMID:10524200. Available from <http://www.ncbi.nlm.nih.gov/pubmed/10524200>
254. Al-Ahmadi W, Al-Ghamdi M, Al-Haj L, Al-Saif M, Khabar KSA. Alternative polyadenylation variants of the RNA binding protein, HuR: Abundance, role of AU-rich elements and auto-regulation. *Nucleic Acids Res* 2009; 37:3612-24; PMID:19359363; <https://doi.org/10.1093/nar/gkp223>
255. Dickson JR, Kruse C, Montagna DR, Finsen B, Wolfe MS. Alternative Polyadenylation and miR-34 family members regulate Tau expression. *J Neurochem* 2013; 1-11
256. Gruber AR, Martin G, Muller P, Schmidt A, Gruber AJ, Gumienny R, Mittal N, Jayachandran R, Pieters J, Keller W, et al. Global 3' UTR shortening has a limited effect on protein abundance in proliferating T cells. *Nat Commun* 2014; 5:5465; PMID:25413384; <https://doi.org/10.1038/ncomms6465>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/25413384>
257. 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; 23:2078-90; PMID:24072873; <https://doi.org/10.1101/gr.156919.113>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/24072873>
258. Spangenberg L, Shigunov P, Abud APR, Cofré AR, Stimamiglio MA, Kuligovski C, Zych J, Schittini A V, Costa ADT, Rebelatto CK, et al. Polysome profiling shows extensive posttranscriptional regulation during human adipocyte stem cell differentiation into adipocytes. *Stem Cell Res* 2013; 11:902-12; PMID:23845413; <https://doi.org/10.1016/j.scr.2013.06.002>
259. Neve J, Furger A. Alternative polyadenylation: Less than meets the eye?. *Biochem Soc Trans* 2014; 42:1190-5; PMID:25110024; <https://doi.org/10.1042/BST20140054>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/25110024>
260. 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; 156:812-24; PMID:24529382; <https://doi.org/10.1016/j.cell.2013.12.026>. Available from <http://www.sciencedirect.com/science/article/pii/S009286741301595X>
261. Gupta I, Clauder-Münster S, Klaus B, Järvelin AI, Aiyar RS, Benes V, Wilkening S, Huber W, Pelechano V, Steinmetz LM. Alternative polyadenylation diversifies post-transcriptional regulation by selective RNA-protein interactions. *Mol Syst Biol* 2014; 10:719; PMID:24569168; <https://doi.org/10.1002/msb.135068>
262. Tranter M, Helsley RN, Paulding WR, McGuinness M, Brokamp C, Haar L, Liu Y, Ren X, Jones WK. Coordinated post-transcriptional regulation of Hsp70.3 gene expression by microRNA and alternative polyadenylation. *J Biol Chem* 2011; 286:29828-37; PMID:21757701; <https://doi.org/10.1074/jbc.M111.221796>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/21757701>
263. Whyteside AR, Turner AJ, Lambert DW. Endothelin-converting enzyme-1 (ECE-1) is post-transcriptionally regulated by alternative polyadenylation. *PLoS One* 2014; 9:e83260; PMID:24497914; <https://doi.org/10.1371/journal.pone.0083260>. Available from <https://doi.org/10.1371%252Fjournal.pone.0083260>
264. Yoon Y, McKenna MC, Rollins DA, Song M, Nuriel T, Gross SS, Xu G, Glatt CE. Anxiety-associated alternative polyadenylation of the serotonin transporter mRNA confers translational regulation by hnRNPK. *Proc Natl Acad Sci U S A* 2013; 110:11624-9; PMID:23798440; <https://doi.org/10.1073/pnas.1301485110>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/23798440>
265. Andreassi C, Riccio A. To localize or not to localize: mRNA fate in 3' UTR ends. *Trends Cell Biol* 2009; 19:465-74; PMID:19716303; <https://doi.org/10.1016/j.tcb.2009.06.001>
266. An JJ, Gharami K, Liao G-Y, Woo NH, Lau AG, Vanevski F, Torre ER, Jones KR, Feng Y, Lu B, et al. Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell* 2008; 134:175-87; PMID:18614020; <https://doi.org/10.1016/j.cell.2008.05.045>
267. Yudin D, Hanz S, Yoo S, Iavnilovitch E, Willis D, Gradus T, Vuppalanchi D, Segal-Ruder Y, Ben-Yaakov K, Hieda M, et al. Localized regulation of axonal RanGTPase controls retrograde injury signaling in peripheral nerve. *Neuron* 2008; 59:241-52; PMID:18667152; <https://doi.org/10.1016/j.neuron.2008.05.029>
268. Andreassi C, Zimmermann C, Mitter R, Fusco S, De Vita S, Saiardi A, Riccio A. An NGF-responsive element targets myo-inositol monophosphatase-1 mRNA to sympathetic neuron axons. *Nat Neurosci* 2010; 13:291-301; PMID:20118926; <https://doi.org/10.1038/nn.2486>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/20118926>
269. Feltrin D, Fusco L, Witte H, Moretti F, Martin K, Letzelter M, Fluri E, Scheiffele P, Pertz O. Growth Cone *MKK7* mRNA Targeting regulates MAP1b-Dependent microtubule bundling to control neurite elongation. *PLoS Biol* 2012; 10:e1001439; PMID:23226105; <https://doi.org/10.1371/journal.pbio.1001439>. Available from <https://doi.org/10.1371%252Fjournal.pbio.1001439>
270. Perry RB-T, Doron-Mandel E, Iavnilovitch E, Rishal I, Dagan SY, Tsoory M, Coppola G, McDonald MK, Gomes C, Geschwind DH, et al. Subcellular knockout of Importin  $\beta$ 1 Perturbs axonal retrograde signaling. *Neuron* 2012; 75:294-305
271. Rhinn H, Qiang L, Yamashita T, Rhee D, Zolin A, Vanti W, Abeliovich A. Alternative  $\alpha$ -synuclein transcript usage as a convergent mechanism in Parkinson's disease pathology. *Nat Commun* 2012; 3:1084. Available from <https://doi.org/10.1038/ncomms2032>
272. Chen L, Decerbo JN, Carmichael GG. Alu element-mediated gene silencing. *EMBO J* 2008; 27:1694-705; PMID:18497743; <https://doi.org/10.1038/emboj.2008.94>
273. Evsyukova I, Bradrick SS, Gregory SG, Garcia-Blanco MA. Cleavage and polyadenylation specificity factor 1 (CPSF1) regulates alternative splicing of interleukin 7 receptor (IL7R) exon 6. *RNA* 2013; 19:103-15; PMID:23151878; <https://doi.org/10.1261/rna.035410.112>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/23151878>
274. Alt FW, Bothwell AL, Knapp M, Siden E, Mather E, Koshland M, Baltimore D. Synthesis of secreted and membrane-bound immunoglobulin mu heavy chains is directed by mRNAs that differ at their 3' ends. *Cell* 1980; 20:293-301; PMID:6771018
275. Lejeune F, Maquat LE. Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Curr Opin Cell Biol* 2005; 17:309-15; PMID:15901502; <https://doi.org/10.1016/j.ceb.2005.03.002>

276. Yao P, Potdar AA, Arif A, Ray PS, Mukhopadhyay R, Willard B, Xu Y, Yan J, Saidel GM, Fox PL. Coding region polyadenylation generates a truncated tRNA synthetase that counters translation repression. *Cell* 2012; 149:88-100; PMID:22386318; <https://doi.org/10.1016/j.cell.2012.02.018>. Available from <http://www.ncbi.nlm.nih.gov/pubmed/22386318>
277. Weng L, Li Y, Xie X, Shi Y. Poly(A) code analyses reveal key determinants for tissue-specific mRNA alternative polyadenylation. *RNA* 2016; 22:813-21; PMID:27095026; <https://doi.org/10.1261/rna.055681.115>. Available from <http://rnajournal.cshlp.org/content/22/6/813.abstract>