

Regulation and function of alternative polyadenylation in development and differentiation

Lorenzo Gallicchio ^a, Gonzalo H. Olivares^b, Cameron W. Berry^c, and Margaret T. Fuller^{a,d}

^aDepartment of Developmental Biology, Stanford University School of Medicine, Stanford, USA; ^bEscuela de Kinesiología, Facultad de Medicina y Ciencias de la Salud, Center for Integrative Biology (CIB), Universidad Mayor, Chile and Department of Neuroscience, Faculty of Medicine, Universidad de Chile, Santiago, Chile; ^cStowers Institute for Medical Research, Kansas City, MO, USA; ^dDepartment of Genetics, Stanford University School of Medicine, Stanford, USA

ABSTRACT

Alternative processing of nascent mRNAs is widespread in eukaryotic organisms and greatly impacts the output of gene expression. Specifically, alternative cleavage and polyadenylation (APA) is a co-transcriptional molecular process that switches the polyadenylation site (PAS) at which a nascent mRNA is cleaved, resulting in mRNA isoforms with different 3'UTR length and content. APA can potentially affect mRNA translation efficiency, localization, stability, and mRNA seeded protein–protein interactions. APA naturally occurs during development and cellular differentiation, with around 70% of human genes displaying APA in particular tissues and cell types. For example, neurons tend to express mRNAs with long 3'UTRs due to preferential processing at PASs more distal than other PASs used in other cell types. In addition, changes in APA mark a variety of pathological states, including many types of cancer, in which mRNAs are preferentially cleaved at more proximal PASs, causing expression of mRNA isoforms with short 3'UTRs. Although APA has been widely reported, both the function of APA in development and the mechanisms that regulate the choice of 3'end cut sites in normal and pathogenic conditions are still poorly understood. In this review, we summarize current understanding of how APA is regulated during development and cellular differentiation and how the resulting change in 3'UTR content affects multiple aspects of gene expression. With APA being a widespread phenomenon, the advent of cutting-edge scientific techniques and the pressing need for *in-vivo* studies, there has never been a better time to delve into the intricate mechanisms of alternative cleavage and polyadenylation.

ARTICLE HISTORY

Revised 11 October 2023
Accepted 17 October 2023

KEYWORDS





Alternative polyadenylation; development; differentiation; RNA processing; translation

1. Introduction

Multicellular organisms are composed of several types of organs and tissues, established during embryonic development and in many cases maintained or repaired throughout life by adult stem cell lineages. The progeny of adult stem cells undergoes tightly orchestrated cellular proliferation and then must differentiate to the proper cell types to regenerate tissues and/or replace cell types lost to normal turnover, injury, or infection. Failure to switch cell state from proliferation to the correct terminal differentiation programme(s) can lead to tissue dysmorphism, scarring, failure to replace lost cell types, and possibly cancer. To achieve proper embryonic development and maintain and repair tissues throughout life, cell fate decisions are profoundly regulated at multiple levels of gene expression. Co-transcriptional alternative processing of nascent transcripts, which produces different mRNA isoforms from the same genetic locus, may be especially useful for programming rapid changes in the proteins expressed as cells transition from one state to the next without the need to re-initiate transcription. The two major co-transcriptional mechanisms that generate different mRNA

isoforms are alternative splicing, which results in mRNA isoforms with different exon combinations [1], and alternative cleavage and polyadenylation (APA), which produces mRNA isoforms with different 3'ends due to terminal cleavage of the nascent transcript at different sites [2]. Alternative processing of nascent mRNAs affects the expression of most human genes and has been linked with serious disease states and developmental disorders [3–6]. In this review, we focus on the role(s) and regulation of Alternative cleavage and polyadenylation (APA) in switches in cell state during development and differentiation.

The final step of mRNA synthesis in eukaryotes is cleavage and polyadenylation, which involves enzymatic cleavage of the nascent transcript and the addition of a string of non-templated adenosines known as the PolyA tail. Important exceptions are histone mRNAs and some long non-coding RNAs, which are cleaved but not polyadenylated [7,8]. Cleavage and polyadenylation occur co-transcriptionally, through recognition of specific polyadenylation site (PAS) motifs in the transcript by the cleavage machinery (CM) [9]. The CM is composed of four major protein complexes: CPSF

CONTACT Lorenzo Gallicchio  gallicchiolorenzo@gmail.com  Department of Developmental Biology, Stanford University School of Medicine, Stanford, USA; Gonzalo H. Olivares  gonzalo.olivares@umayor.cl  Escuela de Kinesiología, Facultad de Medicina y Ciencias de la Salud, Center for Integrative Biology (CIB), Universidad Mayor, Chile and Department of Neuroscience, Faculty of Medicine, Universidad de Chile, Santiago, Chile

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

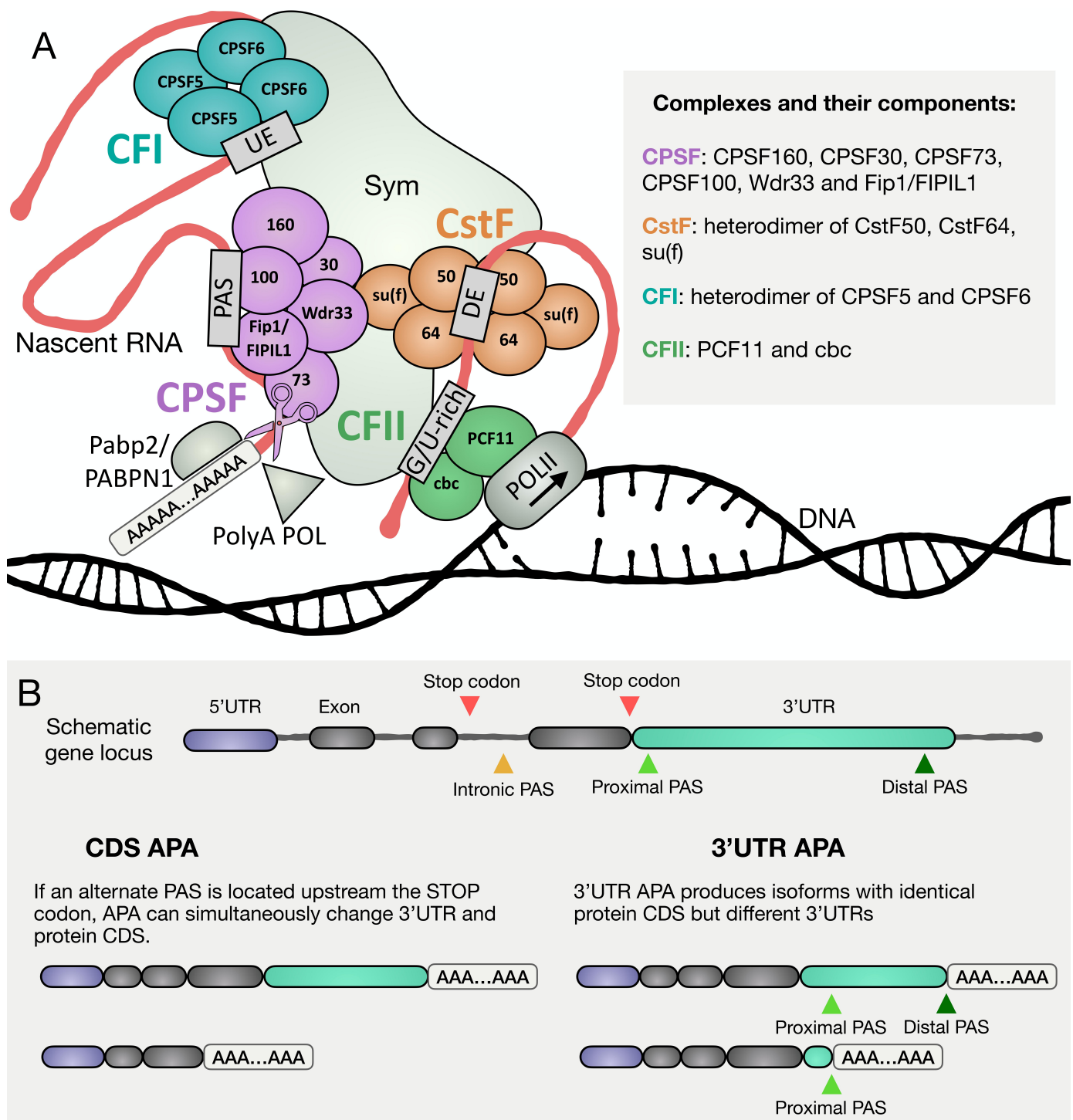


Figure 1. Diagram of 3' end cleavage machinery and outcomes of alternative polyadenylation. (A) scheme representing known cleavage machinery components, showing 4 major protein sub-complexes and additional non-complex proteins. Names of proteins as in *Drosophila melanogaster*. CPSF (in purple) recognizes the PAS and the endonuclease CPSF73 cleaves the nascent RNA molecule. CstF (in orange) is a heterodimer composed of 3 factors that recognizes downstream elements (DE) and enhances the activity of CPSF. CFI (in cyan) recognizes upstream elements (UE), while CFII (in green) is thought to bind G-rich regions and interact through the component PCF11 with the CTD of POLII. Other major components of the CM are the scaffolding protein Symplekin (Sym), the PolyA Polymerase (PolyA POL) and PABP2. (B) alternative polyadenylation (APA) can be divided into coding sequence changing APA (CDS APA) and 3'UTR APA depending on the location of the alternative PASs.

(cleavage and polyadenylation specificity factor), CstF (cleavage stimulation factor), CFI (cleavage factor I), and CFII (cleavage factor II), all knit together by the scaffolding protein Symplekin, plus several other proteins, including nuclear PolyA Polymerase and PolyA binding protein (Figure 1 A).

The CPSF complex recognizes PAS motif (the canonical and most common PAS is 'AAUAAA') and cleaves the nascent RNA 15–30 nucleotides downstream through action of the endonuclease CPSF73. CstF is thought to enhance the activity of CPSF, and CFI modulates cleavage by recognizing specific

upstream elements. CFII is the least characterized cleavage complex; it is thought to bind G-rich regions and interact with the C-terminal domain (CTD) of POLII via the component PCF11 [10,11].

Cleavage and polyadenylation can occur at different sites in a nascent RNA molecule, often signalled by different PAS sequences, resulting in alternate mRNA isoforms. PASs are divided into strong and weak based on the sequence of the PAS itself, as well as additional surrounding motifs that may augment or inhibit binding of CM components or enzymatic processing of the mRNA, once bound. Many genes have proximal PAS motifs that are weaker than their distal PAS motifs [12]. Intriguingly, recent observations indicate that processing of alternate PASs can occur sequentially, with a strong distal PAS better recognized by the CM processed first. The resulting long mRNA can then be retained in the nuclear matrix, possibly by CM components or other undertermined factors, facilitating a second round of processing specified by a weaker proximal PAS, using the retained mRNA as an intermediate to produce a shorter isoform [13].

When a genetic locus has more than one PAS, developmentally regulated or pathologically induced differences in which PAS is chosen for the 3' end cut can change the relative expression levels of the alternate mRNA isoforms in different cell types. Depending on the genetic location alternate PASs occupy on a gene *locus*, APA can change the encoded protein and/or the 3'UTR (Figure 1 B). For example, the use of an alternative PAS in an intronic region upstream of the predominantly used translational stop codon has the potential to simultaneously change both the sequence of the encoded protein as well as the 3'UTR. Such coding region changes can affect cell identity or function due to the expression of alternative or truncated protein isoforms (For a detailed classification of types of APA that could simultaneously change CDS and 3'UTR check this recent review [14]). On the other hand, if the alternatively used PAS sequences are located downstream of the translational stop codon, a switch in usage from one to the other can produce mRNA isoforms that encode the same protein but have different 3'UTRs (termed 3'APA) [15]. As 3'UTRs are hubs of *cis*-regulatory sequences that can regulate translation, stability, and localization of mRNAs by attracting *trans*-acting factors such as microRNAs (miRNAs) and RNA binding proteins (RBPs), developmentally regulated changes in 3'UTRs due to 3'APA can profoundly affect gene product expression by removing or adding functional sequences in the 3'UTR of specific mRNA isoforms. In addition, sequences in the 3'UTR of certain mRNAs have also been shown to alter co-translational folding of the nascent peptide, post-translational modifications of the encoded protein and protein complex assembly [16–19]. Developmentally regulated APA events, therefore, have the potential to profoundly affect both the menu of proteins expressed in specific cell types and the binding partners of nascent proteins.

Since its initial observation in 1980 [20,21], APA has been documented in a variety of organisms, tissues and biological

phenomena and has been linked with severe disease states such as cancer, diabetes and muscular dystrophy [3]. Developmentally regulated use of alternative PAS sites to produce cell type specific mRNA isoforms is common during cell differentiation. For example, for a number of loci preferential cleavage at more distal PAS sites produces mRNA isoforms with longer 3'UTRs in mouse brains and *Drosophila* central nervous system (CNS) than in most other cell types [22,23]. Similarly, 3'UTR lengthening due to APA has been observed in human differentiating cardiomyocytes [24]. Conversely, at a number of loci in differentiating *Drosophila* spermatocytes or mammalian spermatids, cleavage and polyadenylation occur at a more proximal PAS than in earlier male germ cell stages, resulting in mRNA isoforms with truncated 3'UTRs [25–27].

An increasing number of studies in cultured cell lines have investigated the impact of APA on aspects of protein synthesis and mRNA stability and identified factors that can influence PAS choice. However, studies performed *in vivo* remain scarce, resulting in incomplete understanding of the role and regulation of APA during healthy cell differentiation and development. In this review, we aim to present the most recent discoveries on how APA is regulated and how it sets up cell type-specific differential gene product expression, with a focus on cellular differentiation *in vivo*.

2. Regulation of APA

APA has distinct tissue-specific patterns [28,29] and can be regulated by different mechanisms. As detailed below, specific cases of APA have been shown to be influenced by: 1) RNA binding proteins (RBPs) and RNA modifications, 2) expression levels and post-translational modifications of Cleavage Machinery (CM) components, 3) transcriptional dynamics, transcription factor (TF) activity, chromatin conformation, and DNA modifications (Figure 2).

RNA-binding proteins (RBPs) and RNA modifications

Nuclear RNA binding proteins (RBPs) can bind to specific motifs in nascent RNA molecules and positively or negatively direct cleavage by interacting with the transcription machinery or cleavage factors. One of the most famous examples is the *Drosophila* protein ELAV, which in the nervous system binds near proximal PASs of nascent transcripts and promotes POLII readthrough. This results in skipping of the proximal PAS in favour of processing at more distal PASs, leading to expression of mRNA isoforms with long 3'UTRs from target genes [33]. Intriguingly, recruitment of ELAV to the proximal PASs is mediated by POLII pausing, indicating that several modes of APA regulation (RBP recruitment and RNA Polymerase dynamics) could cooperate [34]. ELAV homologs have also been shown to induce expression of mRNA isoforms with long 3'UTRs due to alternative mRNA cleavage in mouse neocortex, where loss of function of the ELAV homolog HuD resulted in many mRNAs shifting to

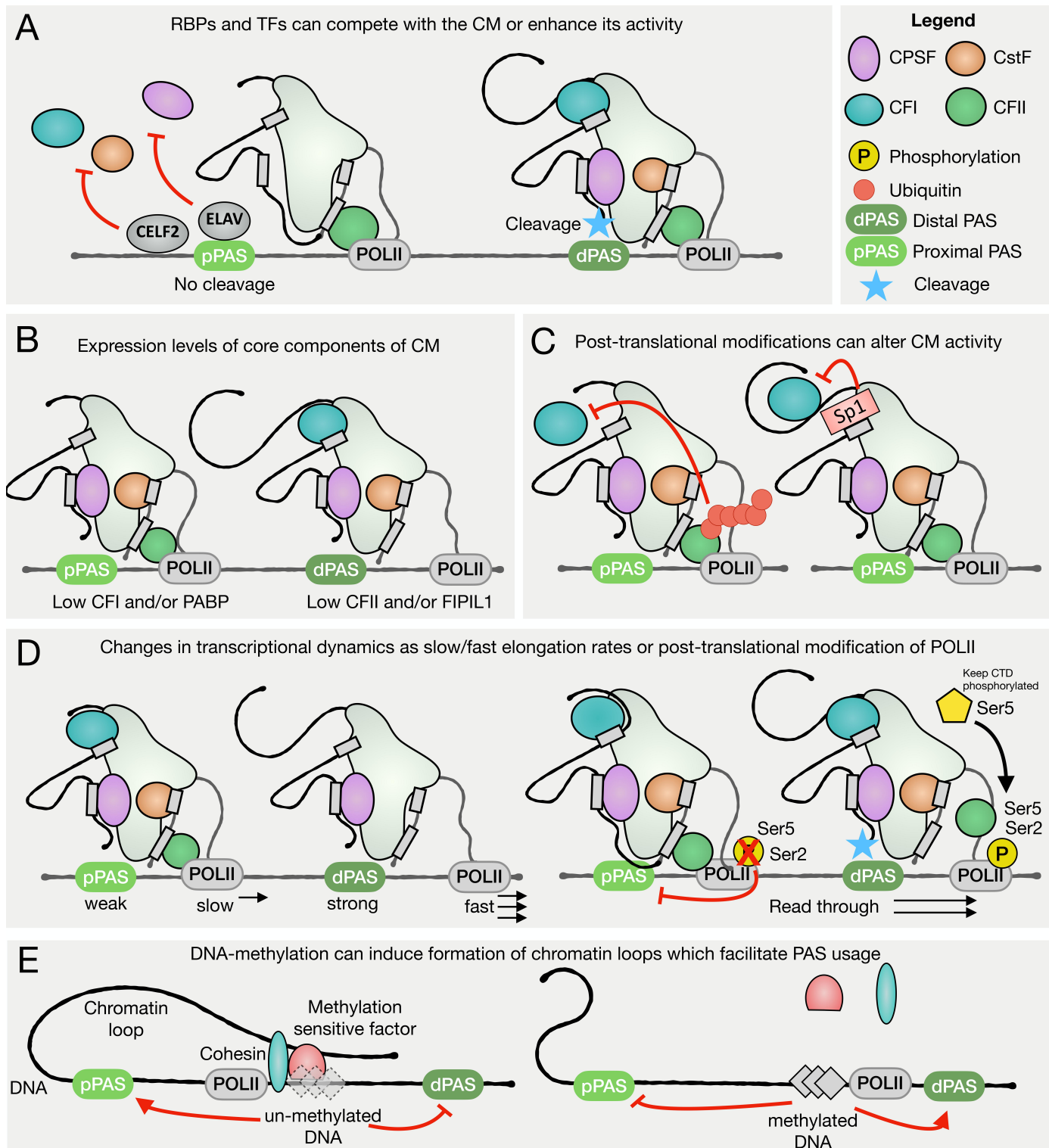


Figure 2. Various mechanisms to regulate APA. (A) RNA binding proteins like ELAV can bind to specific sequences in the proximity of PASs and compete with some of the components of the CM to inhibit cleavage at these specific sites. (B) a change in the expression level of some of the components of the CM can cause preferential processing of specific PASs. For example, down-regulation of CFII has been reported to promote the usage of distal PASs, while down-regulation of CFI can enhance cleavage at proximal PASs. (C) CM activity can be altered by post-translation modification of its core components. For example, ubiquitination of PCF11 by MAGE-A11 or competitive binding of Sp1 can both induce inhibition of CFI activity causing a preferential processing of proximal PASs [30,31]. (D) differential transcription dynamics can induce a change in PAS processing. For instance, while a slowly elongating POLII can augment the recognition on non-canonical weak PAS sites by allowing the correct assembly of the CM, a faster elongating POLII might instead favor strong PASs. In addition, post-translational modification as phosphorylation can result in premature termination of POLII favoring proximal PASs. (E) DNA and RNA modification can alter PAS choice. For example, un-methylated DNA can induce formation of chromatin loops and induce cleavage at more proximal PASs. Figure partially adapted from [32].

cleavage at more proximal sites [35]. Human CELF2, another member of the ELAV-like protein family, has been shown to compete with the CM components CFI and CstF for binding to weak PAS sites. Expression of CELF2 increases in human Jurkat T cells upon stimulation, resulting in APA changes. Interestingly, CELF2 regulates its own transcript APA by repressing the usage of a weak non-canonical PAS site on *CELF2* 3'UTR [36] (Figure 2 A). The *Drosophila* RNA binding protein Sex-Lethal (Sxl) may also in some cases inhibit cleavage of nascent transcripts by competition with components of the CM. Sxl protein binds to multiple sites in the nascent mRNA of the *e(r)* gene in female early germ cells, including a GU-rich element close to the proximal PAS site. This element would also be recognized by the cleavage factor CstF64, binding of which would induce cleavage. However, SXL bound at the site competes with the CM causing skipping of the proximal PAS, inducing alternative processing at the distal site instead [37].

RNA modifications can also alter APA by directing PAS choice. In the developing mouse oocyte, depletion of the nuclear protein YTHDC1, which recognizes the RNA modification N⁶-methyladenosine (m6A), affected alternative splicing of 1966 genes and alternative 3'UTR processing of 864 genes, with a mild trend towards 3'UTR lengthening. YTHDC1 has also been found to interact with CM components [38,39]. Another RNA modification, pseudouridylation, can affect the choice of 3'end cut sites. Downregulation of pseudouridine synthases caused changes in alternative polyadenylation as well as changes in alternative splicing in cultured human cells [40].

Expression levels and post-translational modifications of CM components

Changing expression levels of CM components can affect APA in a cell type and gene specific way, either by promoting usage of distal PAS sites or favouring usage of proximal PAS sites (extensively reviewed here [41]; and here [42]). For example, reduction of CFI components and *PABP* mRNA levels in mouse embryonic fibroblast and in mouse myoblasts resulted in increased usage of more proximal PASs and consequent 3'UTR shortening [43,44]. Consistent with this, in a human cancer-derived cell line, *PABPN1* (*PABP2* in *Drosophila melanogaster*) appeared to prevent cleavage at proximal and weak PAS sites by competing with the CPSF complex, preventing cleavage and therefore inducing 3'UTR lengthening [12]. On the other hand, reducing levels of mRNAs encoding components of CFII in both neuroblastoma and breast cancer cells or *FIPIL1* (*Fip1* in *Drosophila melanogaster*) in mouse myoblasts caused preferential processing at more distal PAS sites, resulting in 3'UTR lengthening [44–46] (Figure 2 B).

Different CM components may regulate APA at different genes, depending on specific *cis*-regulatory sequences located near the alternate PAS sites. For example, CFI appears to increase the usage of a subset of PAS sites in human cells that contain the enhancer motif 'UGUA'. CFI bound at these motifs is thought to interact with the CPSF component *Fip1*, promoting CPSF recruitment to these

PASs and inducing cleavage [47]. Although it has been shown that changing individual CM expression levels can alter APA genome-wide, our understanding of the molecular mechanisms by which each CM component affects APA remains unclear.

Cell type specific changes in the level or action of certain CM components may underlie developmentally regulated APA. During *Drosophila* CNS development, nascent transcripts from several loci are preferentially cleaved at their distal PAS sites, resulting in mRNA isoforms with longer 3'UTRs. Expression of CM transcripts is generally low in *Drosophila* embryos except for the mRNAs encoding the two components of CFI (named in *Drosophila* *CPSF5* and *CPSF6*), which are selectively highly expressed in the CNS. Downregulation by RNAi of CFI component mRNAs in the embryonic CNS changed the site used for 3'end cleavage for many genes, shifting to more distal PAS or more proximal PAS usage, depending on the gene [23]. The observed dual role of CFI (inducing proximal 3'end cleavage for some genes and distal for others) is puzzling and still needs to be fully understood. It might be due to gene-specific sequences (as the observed 'UGUA' element) that may attract CFI to specific sites in a gene dependent manner.

Similarly, changes in level or action of specific CM components during progression towards disease may influence APA associated with pathology. 3'UTR shortening is in fact a common feature observed in many cancer types [48]. The level of mRNA encoding *FIPIL1*, a component of the CPSF complex, has been shown to be the most predictive of 3'UTR shortening in RNAseq datasets derived from leukaemia patients. Indicating a functional relationship, knocking down levels of *FIPIL1* RNA in leukaemia cells restored the APA profile to normal, suggesting *FIPIL1* (named in *Drosophila* *Fip1*) as a possible therapeutic target for leukaemia [49]. In prostate cancer, transcripts from some genes have been shown to be prematurely cleaved at intronic PAS sites upstream of the last protein coding exon, leading to the production of truncated, pro-oncogenic protein isoforms. Recently, Caggiano *et al.* [50] showed that the treatment of prostate cancer cells with hormonal therapy to inhibit androgen receptor signalling suppressed the characteristic usage of premature intronic PAS sites, shifting 3'end cleavage back to the normal distal PAS sites. Interestingly, prostate cancer cells that had acquired resistance to the hormonal therapy showed higher levels of CPSF160. These results raise the possibility that increased levels of the CPSF component CPSF160 may enable cleavage at the upstream, intronic PAS sites and suggest targeting of CPSF160 as a possible therapeutic strategy worth exploring in future studies [50].

Nuclear Poly(A) polymerases (PAP) can also affect PAS choice. Mammalian cells have several distinct nuclear poly(A) polymerases, and PASs site usage appears to be selectively regulated by specific Poly(A) polymerases, depending on the motifs surrounding the PAS. For example, the PAS sites where processing was downregulated by knock-down of the non-canonical STAR-PAP generally had canonical 'AAUAAA' PAS motifs with an upstream AUA motif element. On the other hand, PAS sites where processing was

downregulated by knock-down of the canonical PAP α showed enrichment of 'GCCA' and G-rich motifs upstream of the PAS [51].

The activity of CM components can also be affected by post-translational modifications of the components themselves, as well as by other protein factors that bind to the CM, potentially influencing APA. For example, PCF11, a component of CFII, can be ubiquitinated by the ubiquitin ligase MAGE-A11. This modification of a CFII component inhibits the activity of one of the two CFI components, CFIm25 (alternative name for the *Drosophila* CPSF5) [52]. In cancer cells, MAGE-A11 is upregulated, resulting in increased ubiquitination of PCF11 and a shift to preferential cleavage at proximal PASs [30], which is consistent with the widely reported 3'UTR shortening upon CFI down-regulation. CFI activity can also be inhibited by the transcription factor Sp1. High levels of Sp1 in breast cancer patients were recently shown to correlate with aberrant proximal PAS usage. Sp1 has been found to interact both with the two subunits of CFI and with nascent mRNA, suggesting that high levels of Sp1 could inhibit CFI function by competitively binding near proximal PASs of target mRNAs [31] (Figure 2 C). In addition, many components of the CM have been found to be validated or putative substrates for SUMOylation, covalent linkage of a SUMO (small ubiquitin-like modifier) protein that can alter the function of the target protein. The endonuclease CPSF73 and the scaffolding protein Symplekin are both SUMOylated, with SUMOylation of Symplekin shown to be necessary for Symplekin function, and inhibition of SUMOylation causes general inhibition of 3'end processing [53]. However, whether developmentally regulated changes in SUMOylation of CM components play a role in directing APA switches in healthy biological processes remains to be investigated.

Transcriptional dynamics, transcription factor (TF) activity, chromatin conformation, and DNA modifications

The dynamics of transcription initiation, elongation and termination have also been reported to influence PAS choice [54]. In *Drosophila*, slower transcription elongation rates have been shown to correlate with the usage of proximal and weaker PAS sites, but only in the body and not in the head, suggesting that different cell types might have different modes of APA tissue-specific regulation [55] (Figure 2 D). The *Drosophila* gene *polo* offers a specific example indicating that transcription dynamics can impact PAS choice [56]. The *polo* gene has two PAS sequences in the region of its 3'UTR. In flies carrying a mutation in the gene encoding the core POLII subunit Rbp215 that leads to slower transcript elongation rates, the ratio of processing at the proximal versus the distal PAS sites of *polo* is 3.5-fold higher than in wild-type flies.

Regulated readthrough into a downstream locus caused by programmed failure to terminate nascent transcripts at the normal PAS can have important roles in development. In the *C. elegans* nervous system, Casein Kinase 1 δ (CK1 δ) inhibits

transcription termination via phosphorylation of SSUP-72, which results in Ser5 de-phosphorylation, inhibiting termination and causing RNA Polymerase to read through from the *unc-44* locus into the *Ankyrin* locus. RNA Polymerase read-through of the *unc-44* PAS is necessary for production of the giant form (>10 Kb) of *Ankyrin* transcript, which is necessary for normal nervous system architecture. One of the possible models is that CK1 δ blocks the phosphorylation of the polymerase CTD, inhibiting binding of certain transcription termination factors (such as the cleavage factor CFII), which in turn would provoke POLII readthrough [57] (Figure 2 D).

The recent advent of long-read RNA sequencing is allowing recovery of information on entire mRNA transcripts, avoiding the RNA fragmentation required for standard RNAseq. Anvar et al., 2018 exploited this ability to correlate transcription initiation site(s) and alternative pre-mRNA splicing with other aspects of co-transcriptional mRNA processing, including PAS choice in a cancer cell line and in human tissues. In a recent study, Alfonso-Gonzales et al. used long-read RNA sequencing to compare mRNA isoforms expressed in *Drosophila* brains versus ovaries and discovered that specific transcription start sites (TSS) could correlate with specific PAS choices. Interestingly, they found that many genes that had both alternative TSS and alternative PAS usage also had CREB-binding protein (CBP) bound at both TSS and PAS, strengthening a model in which the CREB-binding protein (CBP) binds specific promoters and creates a 5'-3' loop, affecting PAS usage [58].

Certain transcription factors may also induce alternative polyadenylation. During heart development in mouse, the transcription factor Nkx2-5 not only occupies transcription start sites of target genes but was also found localized to downstream regions of genes. Nkx2-5 recruits the exonuclease Xrn2 to these regions, whose function is to stop transcription. Knock-down of either Nkx2-5 or Xrn2 resulted in expression of mRNA isoforms with longer 3'UTRs, indicating that Nkx2-5 can either directly or indirectly influence APA [59].

DNA methylation, chromatin architecture and *cis*-regulatory DNA sequences can also play roles in regulating APA. Recent results from a human cancer cell line indicated that DNA methylation between two PASs can increase processing at the distal PAS. It is thought that the assembly of a cohesin protein complexed with a methylation-sensitive insulator protein at a site between the two PAS sequences normally induces pausing of RNA polymerase II, leading to 3'end processing at the proximal PAS (Figure 2 E). Methylation of DNA in that region may prevent the assembly of the cohesin complex, allowing transcript extension to the distal site [32]. Chromatin conformation appears to have a role in PAS choice in plants as well. The rice gene *OsASI1* is required for fertility and regulates the timing of flowering. Plants mutant for *OsASI1* show dysregulation of APA at more than 500 genes, switching either from distal PAS to proximal PAS or vice-versa (a higher percentage of genes switched to a more distal PAS). The *OsASI1* protein can bind directly to the intronic heterochromatin of the gene *OsXRNL* and promote production of full length *OsXRNL* transcripts. Interestingly, the class of genes

where APA was most affected by *OsAS11* knockout contained heterochromatic elements, suggesting a link between APA regulation in rice with chromatin conformation [60].

Specific enhancers can also regulate which PAS of a target gene is utilized for 3' end processing, as in the case of the human *PTEN* gene. Processing of nascent *PTEN* transcripts at a proximal and weak PAS was impaired when the *PTEN* enhancer was deleted in breast cancer cells. Additionally, the enhancer-mediated effect on APA is promoter-dependent. In a reporter system, the *PTEN* and *NUDT21* (alternative name for the *Drosophila* *CPSF5*) enhancers induced up-regulation of transcription if the promoter was derived from single 3'UTR genes. However, the same enhancers acting on promoters derived from genes with multiple 3'UTRs induced only changes in cleavage and polyadenylation [61]. Currently, how mechanistically an enhancer could regulate APA still needs to be described.

3. APA regulates mRNA stability, localization, and translation

Alternative Polyadenylation (APA) can lead to the production of mRNA isoforms with different protein coding sequences and/or different 3'UTRs. APA that changes protein coding sequences can result in expression of variant, truncated, aberrantly functional or non-functional proteins (Figure 3 A). On the other hand, APA that changes length and content of 3'UTRs can deeply impact many aspects of mRNA metabolism including: 1) mRNA stability, 2) mRNA localization, 3) mRNA translation and 4) co-translational protein-protein interactions, depending on the 3'UTR sequence included/excluded and the cell type context that determines the *trans*-acting regulators expressed (Figure 3 B).

mRNA stability

In multiple cancer cell lines, APA has been found to produce many transcripts with shorter 3'UTRs compared to corresponding healthy tissue samples. A study investigating the stability of mRNAs in lung, colon, breast, and sarcoma cell lines assessed by Northern blotting following transcriptional inhibition revealed differential stability of mRNA isoforms resulting from APA, with the short 3'UTR isoform being more stable than the long 3'UTR isoform for the three genes tested [62]. However, the genes were pre-selected based on the presence of miRNA binding sites in the long 3'UTR that were cut off in the short 3'UTR isoform. It remains an open question whether these findings hold true globally for the large number of genes that produce mRNAs with short isoforms in cancer cells due to APA.

In human neurons, the *HuR* locus produces a predominant mRNA isoform with a 4.9 kb 3'UTR due to APA, much longer than the ubiquitously-expressed *HuR* isoform, which has a 1.3 kb 3'UTR [63]. Following inhibition of transcription in differentiated P19 neuronal cells, the half-life of the long 3'UTR *HuR* mRNA isoform was measured to be 4–5 hours while the short 3'UTR *HuR*

isoform was much more stable, with a half-life of over 8 hours. The *trans*-acting factors that mediate the more rapid turnover of the long 3'UTR isoform have yet to be identified. Whether expression of the many other mRNAs with longer 3'UTRs in neurons due to APA also leads to a decrease in mRNA stability remains an open question [63].

mRNA localization

The production of mRNAs with different 3'UTRs due to APA can influence the subcellular localization of specific transcript isoforms, likely due to *cis*-regulatory elements present in the extended 3'UTRs. In mice, the *Bdnf* locus produces mRNA isoforms in neurons with either short or long 3'UTRs due to APA. An *et al.* [16] showed that the long *Bdnf* 3'UTR was sufficient to localize a GFP reporter mRNA to neuronal dendrites, while mRNA from a GFP reporter with the short *Bdnf* 3'UTR remained in the cell body [16]. This suggests that isoforms produced by APA can differ in mRNA localization due to *cis*-regulatory elements present in extended 3'UTRs. In a more global study, Ciolli Mattioli *et al.* [64] differentiated excitatory neurons from mESCs and then fractionated the neurons to separate the cell body (soma) from cellular extensions (neurites, including both axons and dendrites). Strikingly, for 522 genes that produced multiple mRNA isoforms in the neurons due to APA, the isoforms are localized differentially to soma vs. neurites ($\log_2\text{FC}$ neurites/soma >1). For 90% of these genes, the shorter 3'UTR isoform was enriched in neurites [64]. The *trans*-acting factors that lead to the differential localization of specific mRNA isoforms remain to be identified.

mRNA translation

A change in the sequences present in 3'UTRs due to APA has the potential to profoundly affect the translation state of mRNA isoforms encoded by the same gene. Consistent with this, evidence for dramatic effects on cell state-specific protein expression has recently begun to appear. After T cell activation, hundreds of transcripts undergo APA to produce mRNA isoforms with shorter 3'UTRs. For five candidate genes subject to such activation-triggered APA, reporter plasmids encoding Luciferase fused to either the full length 3'UTRs or the short common 3'UTR regions were introduced into T cells. Luciferase assays revealed that the longer 3'UTR region led to decreased protein expression, suggesting 3'UTR shortening by APA as a mechanism to up regulate expression of certain proteins upon T cell activation [65]. One of the genes that undergoes 3'UTR shortening due to APA upon T cell activation encodes CD5, which regulates T cell receptor (TCR) signalling. Higher T cell activation resulted in increased levels of CD5 protein expressed in the cells. Shortening of the CD5 3'UTR upon T cell activation removed two miR-204/211 binding sites. Luciferase assays revealed that these were functionally important, acting to repress levels of protein expressed from mRNAs with the long 3'UTR.

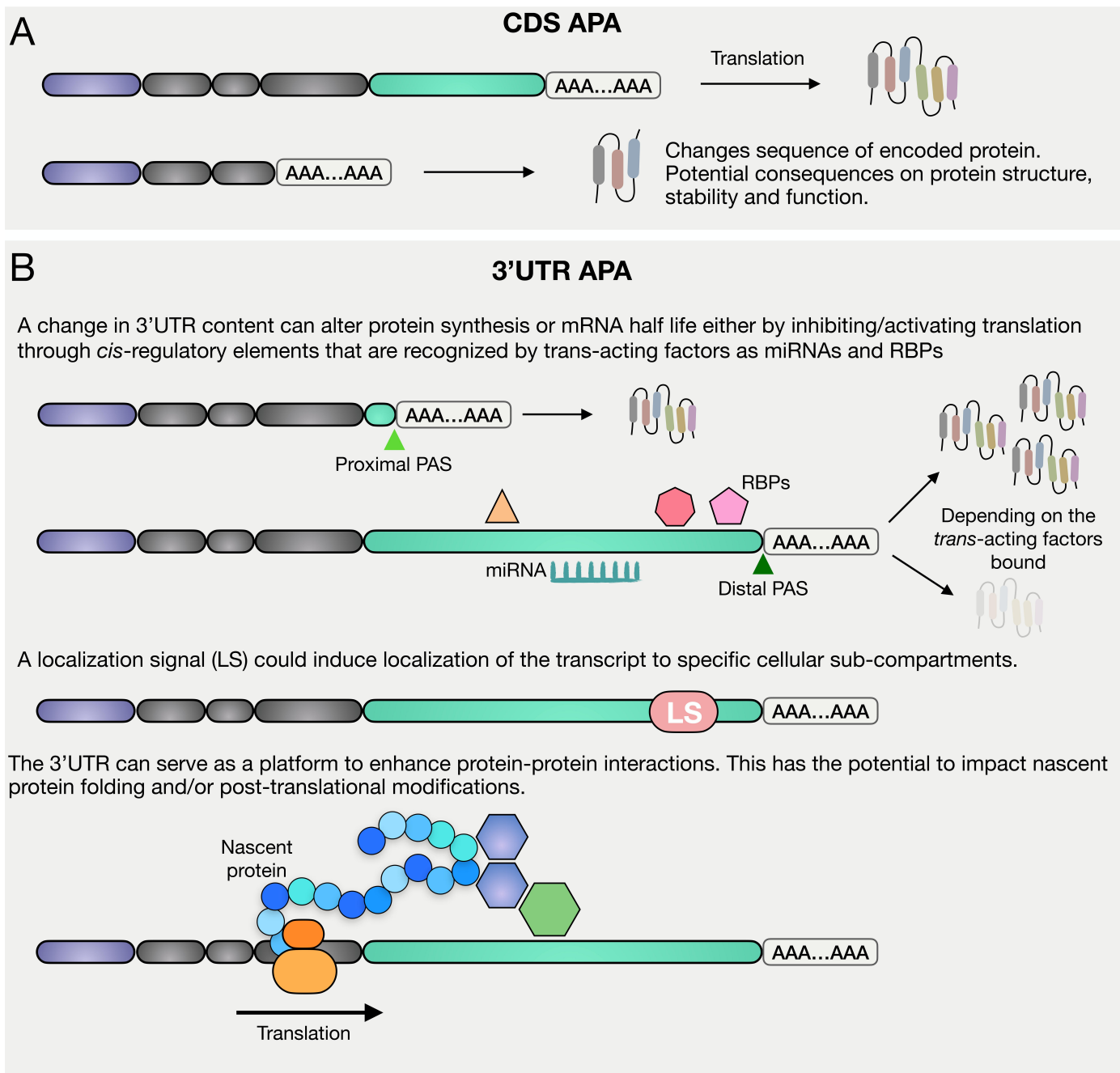


Figure 3. APA can affect protein synthesis in multiple ways. CDS APA can result in the production of multiple mRNA isoforms that can differ in both terminal exon and 3'UTR length and content, causing the possible production of truncated and/or non-functional proteins. 3'UTR APA results in the production of mRNA isoforms with identical CDS but different 3'UTRs. A change in 3'UTR content can cause a switch in translation state, a change in localization of the mRNA molecule, as well as changes in protein-protein interactions.

Luciferase constructs with the long 3'UTR showed lower protein expression than Luciferase constructs with the shortest 3'UTR, but mutation of the two miR-204/211 sites in the long 3'UTR led to higher protein expression [66]. It is notable, however, that global quantification of protein levels in activated vs quiescent T cells by mass spectrometry indicated no overall correlation between abundance of particular proteins at the two stages and the ratio of long 3'UTR vs short 3'UTR isoforms expressed from the corresponding genes, assessed for cases where exactly two alternative PASs were utilized in the cells [67]. This lack of overall correlation may be because

3'UTR extensions cleaved off by APA when T cells are activated can contain either translation-enhancing or translation-repressing sequence elements, depending on the gene.

Developmentally regulated, stage-specific APA is a feature of male germ cell differentiation in animals from *Drosophila* to mammals. As proliferating spermatogonia transition to differentiating spermatocytes in *Drosophila* or as spermatocytes become spermatids in mice widespread stage-specific APA results in expression of mRNA isoforms with shorter 3'UTRs for a select set of genes [25,68,69]. Recently, polysome profiling followed by 3'RNA-Seq revealed that for over half of

the ~500 genes subject to APA leading to 3'UTR shortening in *Drosophila* spermatocytes, the mRNA isoforms with short 3'UTRs migrated differently with respect to polysomes than their counterpart isoforms with long 3'UTRs. For 200 genes the long 3'UTR mRNA isoform expressed in spermatogonia migrated with polysomes, while the short 3'UTR isoform expressed in early spermatocytes migrated with sub-ribosomal fractions. For another 50 genes, the long 3'UTR isoform expressed in spermatogonia migrated in sub-ribosomal fractions, while the short 3'UTR isoform expressed in spermatocytes migrated with ribosomal or polysomal fractions. These results suggest dramatic changes in translation state. Consistent with this, immunofluorescence staining revealed abrupt switches in protein levels, from present in spermatogonia to not detected in spermatocytes or vice versa [25]. Thus, a developmentally regulated switch in the site chosen for making the 3' end cut that terminates particular mRNAs can lead to dramatic changes in the suite of proteins expressed at subsequent stages of cellular differentiation. This can be a switch from protein ON to OFF or protein OFF to ON, presumably depending on the *cis*-acting sequences present in the long 3'UTR expressed in spermatogonia but cleaved off in spermatocytes. For the 200 genes for which polysome fractionation suggested ON to OFF behaviour, the 3'UTR extensions were enriched for PABP binding motifs, which could promote translation of the long 3'UTR transcripts but decreased protein production after removal of the sequences by APA [25]. Further functional work is necessary to identify the precise translational activators and repressors that alter polysome association and protein production from the long and short 3'UTR mRNA isoforms produced by stage-specific APA.

The production of mRNAs with long 3'UTRs due to developmentally regulated APA is a conserved feature of brain tissue. In addition to subcellular localization, the longer 3'UTRs expressed in the brain can also affect translation. Luciferase reporters containing either the short or long 3'UTR isoform expressed in primary hippocampal neurons revealed that the reporter with a short *Bdnf* 3'UTR was more highly translated [70]. Intriguingly, when the hippocampal neurons were treated with tetraethylammonium (TEA) to induce neuronal activation, the reporters containing the long *Bdnf* 3'UTR were now much more highly translated relative to the reporters with short *Bdnf* 3'UTR, indicating that the differential translation depended on neuronal cell state. Taken together with the results of An et al. [16] on subcellular localization discussed above, it is tempting to speculate that the long 3'UTR extension of *Bdnf* transcripts expressed in neurons contains sequences that both target the mRNA to neuronal processes and inhibit its translation until the neuron is activated. Such mechanisms could play important roles in activation dependent strengthening of synapses or other neuronal properties related to learning. More recently, a study combining polysome profiling and deep mRNA sequencing of neurons that had been differentiated *in vitro* from human embryonic stem cells revealed that neuronal transcripts with longer 3'UTRs were less likely to be associated with polysomes [71], consistent with repression of translation. However, dependence on neuronal activation was not studied nor was

the relative translational state of mRNA isoforms partners with short 3'UTRs expressed from the same gene due to APA.

In addition to potential effects on stability of individual mRNA isoforms studied by Mayr and Bartel [62], the production of many transcripts with shorter 3'UTRs due to APA in cancer compared with wild-type tissue can also affect translation of the mRNAs. For three genes (*IMP1*, *DICER-1* and *Cyclin D2*) that show 3'UTR shortening due to APA in cancer compared to non-transformed cell lines, luciferase reporter assays comparing the long or short 3'UTR revealed on average 10-fold higher protein expression from constructs containing the short 3'UTR. Furthermore, mutation of predicted miRNA binding sites in the long 3'UTR led to higher expression of the luciferase reporters, indicating that the long 3'UTRs of the three genes tested are translationally repressed via miRNAs [62]. Additionally, overexpression of an mRNA encoding a candidate oncogene, *IMP-1*, with a short 3'UTR showed six-fold increase in oncogenic transformation capability compared to an *IMP-1* mRNA with long 3'UTR. It is possible that such pathological upregulation of oncogenic proteins due to removal of 3'UTR sequences that normally inhibit translation underlies selection during the progression to cancer for mutations that result in the observed widespread 3'UTRs shortening due to aberrant APA.

Co-translational protein–protein interactions

Recruitment of RNA binding proteins to 3'UTRs can also promote protein–protein interactions with the nascent peptide being translated from the mRNA, resulting in differential protein complex formation, localization, and function. Production of mRNA isoforms with short vs long 3'UTRs due to APA can affect these functions by removing or adding binding sites for these factors. CD47 is a transmembrane protein that acts as a 'don't eat me' signal, blocking engulfment by macrophages when it is displayed on the cell surface. Its role is especially important in allowing cancer cells to avoid attack and clearance by immune system cells. Multiple human cell lines express two major mRNA isoforms for CD47 that differ in 3'UTR length (short and long). Experiments with reporters where either the short or the long CD47 3'UTR was coupled to mRNA encoding a fusion protein in which the extracellular domain of CD47 was replaced by GFP revealed that GFP expressed from the mRNA with the long CD47 3'UTR localized to the cell surface, while GFP expressed from the reporter with the short CD47 3'UTR remained in the endoplasmic reticulum (ER). No differences were identified in the localization of the reporter mRNAs, which were both associated with the ER. Further analysis revealed the mechanism: the HuR/ELAVL1 protein recruited to the CD47 long (but not the short) 3'UTR, binds the SET protein and promotes interaction of SET with a cytoplasmic facing domain of the nascent CD47 (or CD47-GFP) peptide being expressed from the mRNA, with the interaction with SET required for eventual localization of the completed CD47 (or CD47-GFP) to the cell surface. *CD44*, *TNFRSF13C* and *ITGA1* also undergo APA

and were shown to have long 3'UTR isoforms that promote HuR and SET-dependent protein localization [72].

4. Tissue and cell-type specific regulation and function of APA

One of the most interesting aspects of APA is how it can be regulated by developmental programmes or cell state to change the expression of the proteins encoded by target genes. As we have seen above, a change in the site chosen for making the 3' end cut that terminates a nascent transcript can have profound effects on mRNA half-life, transcript localization, and whether and where the encoded protein(s) are expressed. Different genes are subject to APA in different cell types, and the result can either lead to the gene's 3'UTR shortening or lengthening, depending on the tissue or cell state. This indicates that APA could be regulated by different mechanisms in different cell types. Furthermore, shortening (or lengthening) of 3'UTRs can have different, gene-by-gene effects on protein production in the same cell, depending on the particular *cis*-acting sequences in the longer 3'UTR. In this section, we will review some of what is known about the tissue and cell type specific regulation of APA and its functional consequences on cell state *in vivo*, as well as the large areas ripe for discovery about the role APA plays in cell state change in development and disease.

Neural tissue

APA in the nervous system is well documented and conserved across species from insects to mammals. The nervous system stands out for its selective expression of longer 3'UTR isoforms through APA, setting it apart from other tissues, such as the germline (Figure 4 A). In this section, we will explore the fascinating landscape of APA regulation in neural tissue and its functional significance during development and differentiation.

Early gene-specific studies of mouse brain tissues revealed several examples of APA-mediated lengthening of 3'UTRs that hold implications for neural development and function. Comparison of mRNAs expressed in mouse brain vs. other tissues revealed genes that express distinct mRNA isoforms with different 3'UTRs determined by APA. For example, the longest transcript of *Mecp2* (*methyl-CpG-binding protein 2*) is 10.2 kb and is predominant in the brain, while shorter variants (1.8 kb) are found in the liver and lung. *Mecp2* expression in the brain shows a temporal profile, with expression of the long transcript highest during embryonic development, lowest at 12 weeks, and elevated at advanced ages [73]. Notably, levels of MeCP2 protein increase with age in neurons [74], raising the possibility that the long 3'UTR of *Mecp2* may harbour elements that favour translation. Another example of 3'UTR lengthening in the nervous system involves the gene encoding Adducin, an actin capping protein. A brain-specific β -adducin transcript, present in mouse, rat, and human, has an unusually long 3'UTR (up to 6.6 kb) due to tissue-specific alternative polyadenylation. This longer transcript plays a role in synaptic plasticity underlying learning and memory [75].

Data from RNA sequencing in *Drosophila* revealed widespread APA, with a trend towards 3'UTR lengthening in the central nervous system (CNS) [76–78]. Transcripts from around 1000 genes were cleaved at more distal polyadenylation sites in the CNS compared to testis and ovary. It is known that testes show extensive 3' end shortening, however, a fraction (>200) of the genes that show 3' end lengthening in heads vs. testes, showed the same trend in heads vs. all carcasses, strengthening the notion that the CNS tends to express transcripts with long 3'UTRs. The neural transcriptome showed a strong preference for novel distal poly(A) sites, and many highly extended 3'UTRs (up to 18 kb) were previously unannotated, increasing the potential for transcript regulation [77,78]. Genes encoding RNA-binding proteins (RBPs) and transcription factors were preferentially subject to 3'UTR extensions in the CNS. Motif analysis identified enrichment of Pumilio binding sites, U-rich sequences, and binding sites of conserved miRNAs in the neural extensions [78]. As discussed above, ELAV, an RNA-binding protein specific to neurons in *Drosophila*, plays a crucial role in generating mRNA isoforms with 3'UTR extensions during neuronal differentiation. It achieves this by binding to the proximal PAS of target genes, which inhibits 3' end formation near those sites. Consequently, cleavage occurs at downstream PAS sites, leading to the expression of mRNA isoforms with the characteristic neuronal 3'UTR extensions [33].

Systematic analysis of human and mouse tissues revealed increased use of distal polyadenylation sites and expression of *trans*-acting factors involved in 3'UTR processing and polyadenylation [27,79]. Consistently, differentiation of mouse embryonic stem (ES) cells into neurons was accompanied by an increase in the use of distal poly(A) sites, resulting in 3'UTR lengthening [80]. Similarly, zebrafish brains exhibited the longest 3'UTRs, with lower expression of 3'UTR processing factors compared to gonads [81]. The long 3'UTR extensions expressed due to APA in human and mouse neural tissues showed enrichment for binding sites for miRNAs expressed in the brain [22]. However, systematic studies on the global consequences of APA on protein translation in the nervous system are still needed.

The usage of distal PASs resulting in extended 3'UTRs influences mRNA localization in the nervous system. In rat hippocampus, transcripts localized in axons (i.e. neuronal projections) or neuronal cell bodies (somata) exhibited different 3'UTR lengths, with mRNAs localized to the neuropil (composed by axons, dendrites, synapses, glial cell processes and microvasculature) having the longest 3'UTRs reported [82]. As discussed above, the *Brain-derived neurotrophic factor* (*Bdnf*) gene produces mRNA isoforms with short and long 3'UTR through APA, with the long isoforms specifically localized in dendrites of hippocampal neurons [16]. Neuronal activation stimulated localization of long *Bdnf* 3'UTR mRNA to dendrites, playing a crucial role in dendritic functioning [70,83]. Similarly, *RAN-binding protein 1* (*RanBP1*) mRNA isoforms with long 3'UTRs localized to axons, facilitating local translation [84]. In contrast, the *Calmodulin 1* (*Calml1*) gene generated mRNA isoforms with short and long 3'UTRs, with the long isoform restricted to the cell body of dorsal root

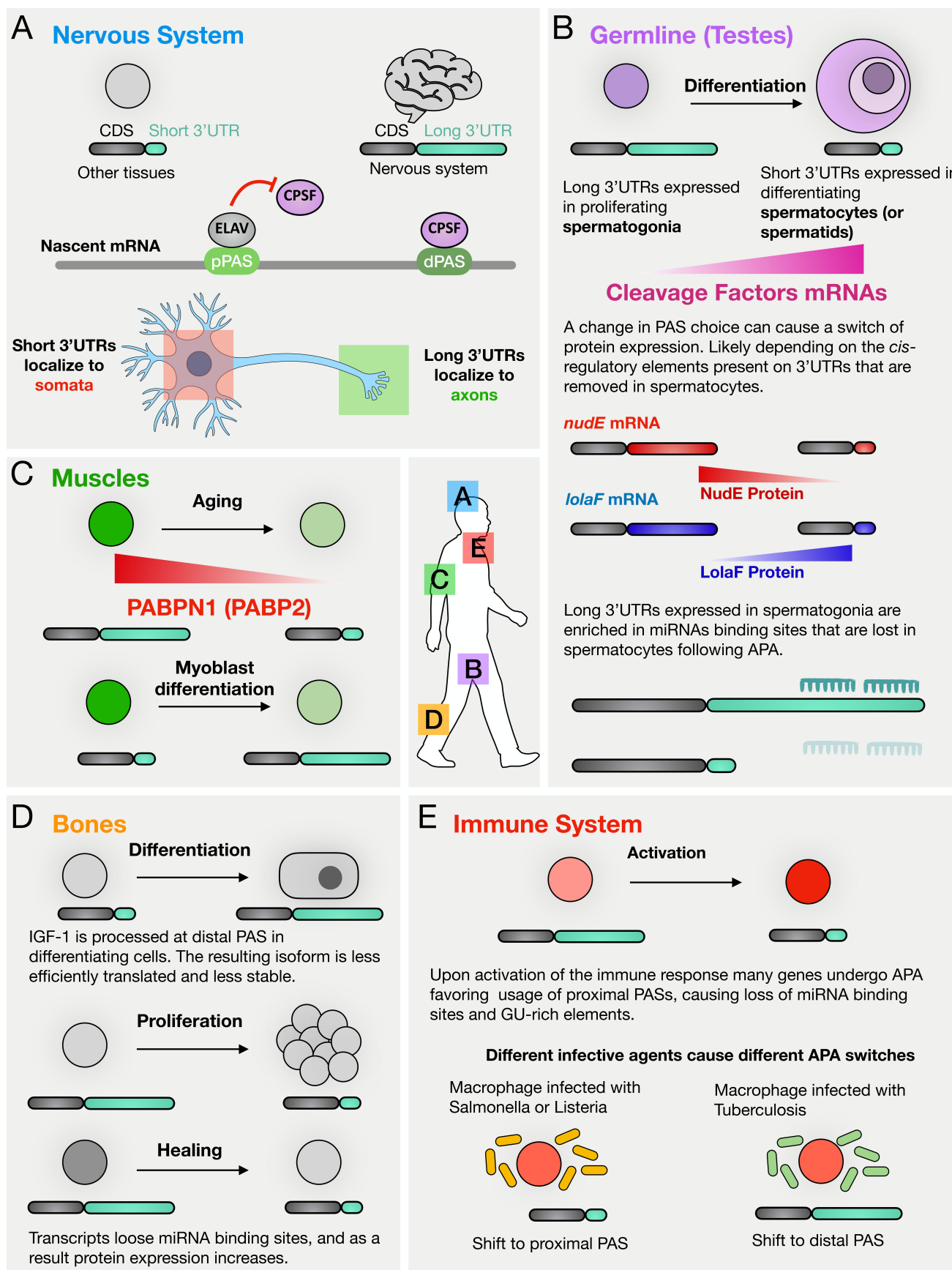


Figure 4. APA has been observed in various tissues and cell-types and its regulation and function are context dependent. (A) in the nervous system APA has been found to mostly induce usage of distal PASs compared to other tissues, resulting in transcripts with particularly long 3'UTRs. For some of the cases studies, such change in 3'UTR length induced transcripts to be localized in axons. (B) in the male germline it has been widely observed that transcripts tend to undergo 3'UTR shortening as spermatogonia differentiate into spermatocytes. A possible mechanism could be the fact that cleavage factors mRNA levels go generally up in spermatocytes compared to spermatogonia. In *Drosophila* this switch of 3'UTR length can affect protein synthesis either from high expression in spermatogonia to

ganglion (DRG) neurons [85]. *Imp1* (myo-inositol monophosphatase-1), an mRNA isoform with a long 3'UTR, was abundant in axons due to the presence of an NGF (nerve growth factor)-dependent axonal localization motif in the extended 3'UTR sequence. Interestingly, axon integrity also required a stable and polyadenylated shorter isoform of the *Imp1* mRNA, which is generated through the direct cytoplasmic cleavage of the longer 3'UTR isoform within axons [86]. Notably, elevated neural activity led to global changes in the expression levels of 3'UTR isoforms and caused significant 3'UTR shortening for 3'UTR isoforms expressed in both neuropil and somata, with a greater effect for those localizing in neuropil [82]. In contrast, although *Tp53inp2* is the predominant transcript with a long 3'UTR localized in axons, it does not undergo translation in sympathetic neurons. Instead, it regulates axon growth in a coding-independent manner through its long 3'UTR, which maintains the transcript in a translationally repressed state [87].

Overall, these findings highlight the importance of APA and extended 3'UTRs in modulating mRNA localization and function in neural tissues, providing new insights into neural development and synaptic plasticity.

Germ line

APA has also been extensively observed in male and female germ cells. In contrast to the 3'UTR lengthening due to APA characteristic of neural development, differentiating male germ cells tend to express transcripts with shorter 3'UTRs compared to progenitor cells or to other tissues. Initially, the analysis of expressed sequence tag (EST) libraries suggested a global trend of 3'UTR shortening in mouse testicular cells compared to somatic cells [26]. During the first wave of spermatogenesis in mice, widespread 3'UTR shortening occurs due to APA, with post-meiotic spermatids having the shortest 3'UTRs. This shortening often eliminated putative mRNA destabilizing elements [68]. The shortening of 3'UTRs at many loci is correlated with increased expression of mRNAs encoding components of the Cleavage machinery (CM) as spermatogenesis proceeded [26].

RNA-seq data from *Drosophila* tissues revealed that the testis transcriptome prefers the use of proximal PAS, resulting in testis 3'UTRs being the shortest among all 29 tissues analysed. The long 3'UTRs forms expressed in the *Drosophila* testis were enriched for miRNA binding sites compared to their shorter counterparts from other tissues, indicating potential functional implications of stage-specific APA in male germ cell differentiation process [77,78]. Similarly, the analysis of 10X snRNA-seq data analysis from the Fly Cell Atlas showed progressive 3'UTR shortening during male germ cell differentiation accompanied with higher mRNA expression levels of cleavage factors [29,88],

resembling the pattern observed in mice [68]. Remarkably, mapping of 3'UTR processing in 14 human samples, including the testis, resulted in similar findings [89]. Analysis of snRNA-seq data also showed that the ovary contains many cell types that show a trend towards the use of proximal PAS [29].

Recent studies on APA in *Drosophila* spermatogenesis showed that developmentally regulated APA can have profound effects on the suite of proteins expressed as cells advance from one stage to the next in a differentiation sequence. In-depth studies by 3'end sequencing of *Drosophila* testes enriched for different germ cell stages from spermatogonia to spermatocytes identified over 500 genes that express mRNA isoforms with a long 3'UTR in proliferating spermatogonia but a short 3'UTR in differentiating spermatocytes due to APA [25]. The switch from transcripts with long 3'UTRs to their short 3'UTR isoforms was in several cases associated with changes in expression of the encoded protein, either from off in spermatogonia to on in spermatocytes (e.g. LolaF) or vice versa (e.g. NudE) (Figure 4 B). Moreover, analysis by polysome gradient fractionation revealed that for more than 250 genes, the long 3'UTR mRNA isoforms expressed in spermatogonia migrated differently than the short 3'UTR mRNA isoforms expressed in spermatocytes, indicating stage-specific changes in translation state [25].

Muscle

APA may also play a role in regulating gene expression during myogenesis and in muscle metabolism. As cultured C2C12 myoblasts differentiate to form myotubes *in vitro*, transcripts from several genes undergo cleavage at more distal, strong PAS sites compared to precursor cells, resulting in 3'UTR lengthening of many transcripts during muscle differentiation [90,91]. On the other hand, the gene encoding the copper transporter *ATP7A* undergoes a shift to more proximal PAS usage during myogenic differentiation, resulting in 3'UTR shortening. This shortening is associated with increased *Atp7a* transcript stability in differentiated myotubes compared to proliferating myoblasts [92]. Differences in 3'UTRs due to APA may also serve as markers for specialized muscle cell types. A total of 405 genes showed significant differences in 3'UTR length of mRNA isoforms expressed in fast vs slow myofibers [93]. Among these, 308 produced mRNA isoforms with longer 3'UTRs in fast fibres (quadriceps muscle) compared to slow fibres (soleus muscle), while 93 produced mRNA isoforms with longer 3'UTRs in slow fibres compared to fast fibres. Thirty-two of the 405 transcripts with distinct 3'UTR isoforms had binding sites for miRNAs that show significant differences between the two muscle types. However, whether these miRNAs functionally regulate these transcripts remains to be explored. Notably, the changes in

low expression in spermatocytes (i.e. nudE) or vice-versa (i.e. lolaF). (C) in ageing muscle cells decreasing levels of PABPN1 (in *Drosophila* Pabp2) are correlated with 3'UTR shortening, while during myoblast differentiation transcripts tend to undergo 3'UTR lengthening. (D) in bone tissue, APA can induce lengthening or shortening of 3'UTR depending on the biological process: while during differentiation it has been observed a trend towards processing of distal PAS, during proliferation and healing many transcripts undergo 3'UTR shortening. (E) in the immune system it has been observed that many transcripts are cleaved at proximal PASs after activation of immune response. Interestingly, macrophages infected by different agents show switches of APA which change depending on the agent: while upon infection with salmonella or listeria transcripts tend to be cleaved at proximal PAS, upon tuberculosis infection transcripts tend to be cleaved at more distal PASs.

APA observed in global studies did not correlate with differences in levels of mRNA isoforms between the two muscle types [93].

Satellite cells, adult stem cells responsible for skeletal muscle maintenance and repair following injury, exist in a quiescent state in healthy muscles and can be activated to re-enter the cell cycle and undergo multiple divisions when appropriately stimulated [94]. Both *PAX3* and *PAX7* transcripts encode transcription factors that are key regulators of myogenesis [95]. In quiescent stem cells, these transcripts undergo APA, leading to the generation of mRNA isoforms with either short 3'UTRs (*PAX3*) or alternative C-terminal domains (*PAX7*) [96,97]. In skeletal muscle stem cells, usage of the proximal PAS of the *PAX3* gene is repressed by the U1 small nucleolar RNA, and the resulting long 3'UTR transcripts are translationally regulated by the microRNA miR-206 [98]. However, when U1 levels are low in diaphragm muscle stem cells, *PAX3* transcripts with a short 3'UTR are produced, which evade regulation by the muscle-specific miR-206 [96,98].

Another mechanism that might regulate APA during muscle cell development and homeostasis involves the levels of expression of the nuclear polyA binding protein PABPN1 (in *Drosophila* and in Figure 1 A is Pabp2). Aging muscles showed diminished levels of PABPN1, contributing to muscle wasting and the development of myogenic defects [12,99,100]. Decreased PABPN1 levels led to reduced utilization of distal PAS on specific genes and adversely impacted translation efficiency, thereby compromising protein homeostasis in age-related muscle atrophy [100] (Figure 4 C). Exploring potential therapeutic strategies centred around modulating PABPN1 levels holds promise for future investigations in this area.

Bone

Changes in 3'UTR sequences due to APA have been shown to play important roles during bone formation. Insulin-like growth factor 1 (IGF-1) is essential for skeletal growth, with low levels linked to reduced bone mass in humans and mice. C3H/He/J mice express higher levels of skeletal IGF-1 and have higher bone mass than C57BL/6J mice. During osteoblastic differentiation in C57BL/6J cells, nascent IGF-1 transcripts tend to be cleaved at a more distal PAS resulting in increased expression of an mRNA isoform with a long 3'UTR (approximately 6300 bases rather than the 200 bases expressed in osteoblasts of C3H/He/J mice) [101]. This long 3'UTR mRNA isoform is less stable and has lower translation efficiency than the isoform with a short 3'UTR. This may be because the long *IGF-1* 3'UTR contains [GU]₁₈ repeats that can be targeted by miR-29 and miR-365, both of which increase in expression during osteoblastic differentiation [101].

During endochondral ossification, a key process for bone development and healing, there is widespread APA in proliferating cells between weeks 2 and 3 of bone healing, with a global shift towards proximal PAS usage leading to an increase in the abundance of mRNAs isoforms with short 3'UTRs [102]. For example, cleavage near the proximal PAS

of nascent transcripts from the *collagen type I alpha 1* (*Col1a1*) and *alpha 2* (*Col1a2*) genes was associated with an increase in mRNA levels, partly due to the removal of miR-29a-3p binding sites that target the longer 3'UTR isoforms of the mRNAs for degradation [102]. However, there is limited information regarding the global-scale consequences on protein expression and the underlying mechanisms that regulate APA in response to tissue damage in bone (Figure 4 D).

Immune system

APA plays crucial roles in the immune system, regulating gene expression related to cell proliferation, differentiation, and immune responses in myeloid and lymphoid lineages. APA-mediated regulation of IgM heavy chain isoforms highlights its significant impact on specific protein expression, affecting immune cell function and differentiation processes. During mouse primary B cell differentiation into plasma cells, there is a switch from the membrane-bound (μ m) to the secreted form (μ s) of IgM heavy chain (μ) due to CDS APA [21,103–106]. Expression of CstF64 is repressed in mouse primary B cells, raising the possibility that expression levels of components of the CM may developmentally regulate this switch. Up-regulation of CstF64 expression in B cells triggered the switch from μ m to μ s isoforms [106,107]. Notably, CstF64 binds with higher affinity to the PAS site of the μ m isoform indicating its specificity [107,108]. However, it has been shown that the increase in CstF64 expression associated with the G0 to S transition in human B cells is independent of the switch from μ s to μ m IgM heavy chain isoform, suggesting that CstF64 levels may not be necessary to mediate the differentiation-induced switch [109].

Macrophages are crucial for innate immunity and respond dynamically to infections. Exposure to pathogens such as *Listeria* or *Salmonella* triggers a shift towards proximal PAS usage in several genes, leading to expression of mRNA isoforms with shorter 3'UTRs that lack binding sites for certain immune-activated miRNAs [110]. This may allow the transcripts to evade translational repression by specific miRNAs and promote gene expression during the immune response [110]. Similarly, vesicular stomatitis virus infection induces preferential proximal PAS usage, enhancing the translation of specific genes in both human and mouse macrophages [111]. However, the regulation of APA in macrophages is context dependent, varying with type of infection (Figure 4 E). For example, *Mycobacterium tuberculosis* infection leads to extensive distal PAS usage, resulting in the expression of mRNA isoforms with longer 3'UTRs [112].

Changes in 3'end cleavage site have been observed during cell proliferation in the immune response, particularly in activated murine and human CD4⁺ T lymphocytes. Activation with anti-CD3/anti-CD28 antibodies induced widespread APA, leading to expression of mRNA isoforms with shorter 3'UTRs due to preferential cleavage at proximal PAS sites [65]. Additionally, genes encoding key regulators of T cell signalling such as CD5 and NF-ATc undergo APA leading to 3'UTR shortening upon T cell activation, further supporting the role of APA in modulating gene expression during immune responses. The resulting 3'UTR shortening

often removes miRNA binding sites present in long 3'UTR isoforms, leading to increased protein expression from 3'UTR reporters [65,66,113]. However, global analysis showed that the observed 3'UTR shortening in proliferating T cells was not always associated with changes in mRNA or protein levels, as might be expected due to the loss of miRNAs target sites and destabilizing sequence elements [67].

5. Conclusions

Although APA was first observed 4 decades ago, only recently we begun to understand the impact and widespread use of this co-transcriptional molecular mechanism. It is becoming evident that modifying the menu of transcribed genes at a given time is not the only strategy a cell uses to drive changes in the suite of proteins expressed. Alternative processing of mRNAs causes the production of different mRNA isoforms from the same genetic locus, potentially changing the type and/or expression levels of the encoded protein. Specifically, alternative polyadenylation (APA) has been now observed in a considerable number of biological phenomena, including during normal cellular differentiation and in many disease states. Indeed, more than 70% of human genes are subject to alternative polyadenylation in one tissue or another. It is therefore extremely important to deepen our understanding of how APA is regulated, how it modulates gene expression during healthy differentiation and developmental processes, and how APA mis-regulation leads to serious disease states. Clinically, APA can be used as a biomarker for some kinds of cancers [114], and studies that aim to understand the role of APA in cancer cells have the potential to identify new therapeutic targets. For example, analysis of APA in pancreatic ductal adenocarcinomas datasets identified the casein kinase CSNK1A1 as an APA-regulated target, with abnormal 3'UTR shortening leading to increased expression in cancer cells. Knocking down CSNK1A1 in cancer cells diminished cell proliferation as well as cancer growth [115].

Recent years have seen the development of a great number of sequencing techniques and bioinformatic tools (for a recent review see [116]: that are proving to be extremely helpful in deepening our understanding of APA regulation and function). For example, FLAM-seq, a recently developed technique for sequencing full-length polyA tails, showed that APA can influence PolyA tail length, pointing towards an additional mechanism of regulation by APA [117]. Studies employing single-cell resolution (e.g. single-cell or single nuclear RNA seq) have added detailed information on changes in APA during development and differentiation across *Drosophila* tissues [29] and during mouse embryogenesis [28]. Single-cell sequencing has the power to capture not only the different cell types within a given tissue but also specific developmental snapshots of cells undergoing differentiation. An interesting recent study performed single-cell polyadenylation sequencing in many cell lines and found that genes that show multiple PASs in bulk RNAseq data, show a single PAS that changes cell by cell, indication cell-to-cell variability of APA [118]. Many single-cell studies investigating APA have been performed in cancer cells [119,120]. However, we still lack single-cell resolution investigation of APA profiles and regulation during specific developmental processes and differentiation programmes *in vivo*. APA has now been observed in many biological contexts and processes, thanks

to fast developing experimental and computational tools, and APA has been demonstrated to regulate gene expression on a gene-by-gene basis, either by affecting mRNA stability, localization, or translation. The next great challenge is to employ genome-wide and high throughput techniques to answer more systematically basic questions about how APA functions in developmental and cellular differentiation. We suggest that APA could serve as a co-transcriptional mechanism to facilitate clean and fast switches in the menu of proteins expressed by cells at sequential steps of differentiation and/or development. Future work will elucidate this model.

Looking to the future, it is particularly important to understand how APA is regulated during normal differentiation and development, and which are the tissue- or disease-specific *trans*-acting factors that bind to *cis*-regulatory regions of 3'UTRs removed/added by APA and how they affect protein production and function. Although it has been shown that APA can be regulated by levels of CM components, transcriptional dynamics, DNA modifications, etc. in certain situations, how and in which contexts these possible regulatory mechanisms control the choice of 3' end cut site during normal differentiation and development to drive expression of different mRNA isoforms is not understood. We also do not understand the molecular mechanisms by which processing at specific PASs is affected when APA is mis regulated by modulating CM expression levels, transcriptional dynamics, DNA modifications, etc. In addition, while many gene-specific examples have been reported of miRNAs or RNA binding proteins that inhibit or activate translation when bound to specific 3'UTR sequences, a global analysis of how and at which level APA affects protein synthesis is still lacking for most of the systems in which APA has been observed. Techniques such as polysome profiling and ribosome profiling coupled with next-generation sequencing and the growing number of bioinformatic tools to identify and quantify APA events provide a powerful trail map for future studies aimed to fully characterize APA regulation and function in development and cellular differentiation.

Acknowledgments

We thank the members of the Fuller lab for opinions about this manuscript during the writing process. We acknowledge the valuable contributions of numerous research papers in the field, although, due to space constraints and the focus of this review, it was not feasible to cite them all individually. Their collective efforts have undoubtedly shaped our understanding and informed the context of this work. Structure and design of the manuscript was formulated by L.G., and it was jointly written by L.G., C.W.B. and G.O. L.G. designed and drew the figures. M.T.F. revised and provided valuable comments that shaped the manuscript. This manuscript is dedicated to Leopoldo.

Disclosure statement

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

Funding

The work was supported by an American-Italian Cancer Foundation postdoctoral fellowship (2021-2023) awarded to LG, by NIH grant R35 GM136433 and funds from the Katherine Dexter McCormick and Stanley McCormick Memorial Professorship to MTF, and by a Stanford Medical School Dean's Discovery Innovation Fund Award to MTF.

ORCID

Lorenzo Gallicchio  <http://orcid.org/0000-0002-5040-0139>

References

- [1] Ule J, Blencowe BJ. Alternative splicing regulatory networks: functions, mechanisms, and evolution. *Mol Cell*. 2019;76(2):329–345. doi: [10.1016/j.molcel.2019.09.017](https://doi.org/10.1016/j.molcel.2019.09.017)
- [2] Ren F, Zhang N, Zhang L, et al. Alternative polyadenylation: a new frontier in post transcriptional regulation. *Biomark Res*. 2020;8(1). doi: [10.1186/s40364-020-00249-6](https://doi.org/10.1186/s40364-020-00249-6)
- [3] Gruber AJ, Zavolan M. Alternative cleavage and polyadenylation in health and disease. *Nat Rev Genet*. 2019;20(10):599–614. doi: [10.1038/s41576-019-0145-z](https://doi.org/10.1038/s41576-019-0145-z)
- [4] Montes M, Sanford BL, Comiskey DF, et al. RNA splicing and disease: animal models to therapies. *Trends Genet*. 2019;35(1):68–87. doi: [10.1016/j.tig.2018.10.002](https://doi.org/10.1016/j.tig.2018.10.002)
- [5] Wang X, Li J, Bian X, et al. CircURI1 interacts with hnRNPM to inhibit metastasis by modulating alternative splicing in gastric cancer. *Proc Natl Acad Sci, USA*. 2021;118(33). doi: [10.1073/pnas.2012881118](https://doi.org/10.1073/pnas.2012881118)
- [6] Xue Z, Warren RL, Gibb EA, et al. Recurrent tumor-specific regulation of alternative polyadenylation of cancer-related genes. *BMC Genomics*. 2018;19(1). doi: [10.1186/s12864-018-4903-7](https://doi.org/10.1186/s12864-018-4903-7)
- [7] Marzluff WF, Wagner EJ, Duronio RJ. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet*. 2008;9(11):843–854. doi: [10.1038/nrg2438](https://doi.org/10.1038/nrg2438)
- [8] Yang L, Duff MO, Graveley BR, et al. Genomewide characterization of non-polyadenylated RNAs. *Genome Biol*. 2011;12(2):R16. doi: [10.1186/gb-2011-12-2-r16](https://doi.org/10.1186/gb-2011-12-2-r16)
- [9] Levitt N, Briggs D, Gil A, et al. Definition of an efficient synthetic poly(A) site. *Genes Dev*. 1989;3(7):1019–1025. doi: [10.1101/gad.3.7.1019](https://doi.org/10.1101/gad.3.7.1019)
- [10] Sun Y, Hamilton K, Tong L. Recent molecular insights into canonical pre-mRNA 3'-end processing. *Transcription*. 2020;11(2):83–96. doi: [10.1080/21541264.2020.1777047](https://doi.org/10.1080/21541264.2020.1777047)
- [11] Zhang Z, Gilmour DS. Pcf11 is a termination factor in *Drosophila* that dismantles the elongation complex by bridging the CTD of RNA Polymerase II to the nascent transcript. *Mol Cell*. 2006;21(1):65–74. doi: [10.1016/j.molcel.2005.11.002](https://doi.org/10.1016/j.molcel.2005.11.002)
- [12] Jenal M, Elkon R, Loayza-Puch F, et al. The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell*. 2012;149(3):538–553. doi: [10.1016/j.cell.2012.03.022](https://doi.org/10.1016/j.cell.2012.03.022)
- [13] Tang P, Yang Y, Li G, et al. Alternative polyadenylation by sequential activation of distal and proximal PolyA sites. *Nat Struct Mol Biol*. 2022;29(1):21–31. doi: [10.1038/s41594-021-00709-z](https://doi.org/10.1038/s41594-021-00709-z)
- [14] Zhang Y, Liu L, Qiu Q, et al. Alternative polyadenylation: methods, mechanism, function, and role in cancer. *J Exp Clin Cancer Res*. 2021;40(1):51. doi: [10.1186/s13046-021-01852-7](https://doi.org/10.1186/s13046-021-01852-7)
- [15] Tian B, Manley JL. Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol*. 2017;18(1):18–30. doi: [10.1038/nrm.2016.116](https://doi.org/10.1038/nrm.2016.116)
- [16] An JJ, Gharami K, Liao GY, et al. Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell*. 2008;134(1):175–187. doi: [10.1016/j.cell.2008.05.045](https://doi.org/10.1016/j.cell.2008.05.045)
- [17] Floor SN, Doudna JA. Tunable protein synthesis by transcript isoforms in human cells. *Elife*. 2016;5. doi: [10.7554/eLife.10921](https://doi.org/10.7554/eLife.10921)
- [18] Graham RR, Kyogoku C, Sigurdsson S, et al. Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. *Proc Natl Acad Sci U S A*. 2007;104(16):6758–6763. doi: [10.1073/pnas.0701266104](https://doi.org/10.1073/pnas.0701266104)
- [19] Mayr C. What are 3' utrs doing? *Cold Spring Harb. Perspect Biol*. 2019;11(10):a034728. doi: [10.1101/cshperspect.a034728](https://doi.org/10.1101/cshperspect.a034728)
- [20] Alt FW, Bothwell ALM, Knapp M, et al. Synthesis of secreted and membrane-bound immunoglobulin mu heavy chains is directed by mRNAs that differ at their 3' ends. *Cell*. 1980;20(2):293–301. doi: [10.1016/0092-8674\(80\)90615-7](https://doi.org/10.1016/0092-8674(80)90615-7)
- [21] Early P, Rogers J, Davis M, et al. Two mRNAs can be produced from a single immunoglobulin μ gene by alternative RNA processing pathways. *Cell*. 1980;20(2):313–319. doi: [10.1016/0092-8674\(80\)90617-0](https://doi.org/10.1016/0092-8674(80)90617-0)
- [22] Miura P, Shenker S, Andreu-Agullo C, et al. Widespread and extensive lengthening of 3' UTRs in the mammalian brain. *Genome Res*. 2013. doi: [10.1101/gr.146886.112](https://doi.org/10.1101/gr.146886.112)
- [23] Vallejos Baier R, Picao-Osorio J, Alonso CR. Molecular regulation of alternative polyadenylation (APA) within the *Drosophila* nervous system. *J Mol Biol*. 2017;429(21):3290–3300. doi: [10.1016/j.jmb.2017.03.028](https://doi.org/10.1016/j.jmb.2017.03.028)
- [24] Yang Y, Wu X, Yang W, et al. Dynamic alternative polyadenylation during iPSC differentiation into cardiomyocytes. *Comput Struct Biotechnol J*. 2022;20:5859–5869. doi: [10.1016/j.csbj.2022.10.025](https://doi.org/10.1016/j.csbj.2022.10.025)
- [25] Berry CW, Olivares GH, Gallicchio L, et al. Developmentally regulated alternate 3' end cleavage of nascent transcripts controls dynamic changes in protein expression in an adult stem cell lineage. *Genes Dev*. 2022;36(15–16):916–935. doi: [10.1101/gad.349689.122](https://doi.org/10.1101/gad.349689.122)
- [26] Liu D, Brockman JM, Dass B, et al. Systematic variation in mRNA 3'-processing signals during mouse spermatogenesis. *Nucleic Acids Res*. 2007;35(1):234–246. doi: [10.1093/nar/gkl919](https://doi.org/10.1093/nar/gkl919)
- [27] Zhang H, Lee JY, Tian B. Biased alternative polyadenylation in human tissues. *Genome Biol*. 2005;6(12):R100. doi: [10.1186/gb-2005-6-12-r100](https://doi.org/10.1186/gb-2005-6-12-r100)
- [28] Agarwal V, Lopez-Darwin S, Kelley DR, et al. The landscape of alternative polyadenylation in single cells of the developing mouse embryo. *Nat Commun*. 2021;12(1):5101. doi: [10.1038/s41467-021-25388-8](https://doi.org/10.1038/s41467-021-25388-8)
- [29] Lee S, Chen Y-C, Gillen AE, et al. Diverse cell-specific patterns of alternative polyadenylation in *Drosophila*. *Nat Commun*. 2022;13(1):5372. doi: [10.1038/s41467-022-32305-0](https://doi.org/10.1038/s41467-022-32305-0)
- [30] Yang SW, Li L, Connelly JP, et al. A cancer-specific ubiquitin ligase drives mRNA alternative polyadenylation by ubiquitinating the mRNA 3' end processing complex. *Mol Cell*. 2020;77(6):1206–1221.e7. doi: [10.1016/j.molcel.2019.12.022](https://doi.org/10.1016/j.molcel.2019.12.022)
- [31] Song J, Nabeel-Shah S, Pu S, et al. Regulation of alternative polyadenylation by the C2H2-zinc-finger protein Sp1. *Mol Cell*. 2022;82(17):3135–3150.e9. doi: [10.1016/j.molcel.2022.06.031](https://doi.org/10.1016/j.molcel.2022.06.031)
- [32] Nanavaty V, Abrash EW, Hong C, et al. DNA methylation regulates alternative polyadenylation via CTCF and the cohesin complex. *Mol Cell*. 2020;78(4):752–764.e6. doi: [10.1016/j.molcel.2020.03.024](https://doi.org/10.1016/j.molcel.2020.03.024)
- [33] Hilgers V, Lemke SB, Levine M. ELAV mediates 3' UTR extension in the *Drosophila* nervous system. *Genes Dev*. 2012;26(20):2259–2264. doi: [10.1101/gad.199653.112](https://doi.org/10.1101/gad.199653.112)
- [34] Oktaba K, Zhang W, Lotz TS, et al. ELAV links paused pol II to alternative polyadenylation in the *Drosophila*. *Nervous Sys Mol Cell*. 2015;57:341–348.
- [35] Sena RM, Twiss JL, Gardiner AS, et al. The RNA-Binding protein HuD regulates alternative splicing and alternative polyadenylation in the mouse neocortex. *Molecules*. 2021;26(10):2836. doi: [10.3390/molecules26102836](https://doi.org/10.3390/molecules26102836)
- [36] Chatrikhi R, Mallory MJ, Gazzara MR, et al. RNA binding protein CELF2 regulates signal-induced alternative polyadenylation by competing with enhancers of the polyadenylation machinery. *Cell Re*. 2019;28(11):2795–2806.e3. doi: [10.1016/j.celrep.2019.08.022](https://doi.org/10.1016/j.celrep.2019.08.022)
- [37] Gawande B, Robida MD, Rahn A, et al. *Drosophila* Sex-lethal protein mediates polyadenylation switching in the female germline. *EMBO J*. 2006;25(6):1263–1272. doi: [10.1038/sj.emboj.7601022](https://doi.org/10.1038/sj.emboj.7601022)
- [38] Kasowitz SD, Ma J, Anderson SJ, et al. Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. *PLoS Genet*. 2018;14(5):e1007412. doi: [10.1371/journal.pgen.1007412](https://doi.org/10.1371/journal.pgen.1007412)
- [39] Ke S, Alemu EA, Mertens C, et al. A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev*. 2015;29(19):2037–2053. doi: [10.1101/gad.269415.115](https://doi.org/10.1101/gad.269415.115)
- [40] Martinez NM, Su A, Burns MC, et al. Pseudouridine synthases modify human pre-mRNA co-transcriptionally and affect

- pre-mRNA processing. *Mol Cell*. 2022;82(3):645–659.e9. doi: [10.1016/j.molcel.2021.12.023](https://doi.org/10.1016/j.molcel.2021.12.023)
- [41] Pereira-Castro I, Moreira A. On the function and relevance of alternative 3'-UTRs in gene expression regulation. *WIREs RNA*. 2021;12(5):e1653. doi: [10.1002/wrna.1653](https://doi.org/10.1002/wrna.1653)
- [42] Mitschka S, Mayr C. Context-specific regulation and function of mRNA alternative polyadenylation. *Nat Rev Mol Cell Biol*. 2022;23(12):779–796. doi: [10.1038/s41580-022-00507-5](https://doi.org/10.1038/s41580-022-00507-5)
- [43] Brumbaugh J, Di Stefano B, Wang X, et al. Nudt21 controls cell fate by connecting alternative polyadenylation to chromatin signaling. *Cell*. 2018;172(1–2):106–120.e21. doi: [10.1016/j.cell.2017.11.023](https://doi.org/10.1016/j.cell.2017.11.023)
- [44] Li W, You B, Hoque M, et al. Systematic profiling of poly(A)+ transcripts modulated by core 3' end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. *PLoS Genet*. 2015;11(4):e1005166. doi: [10.1371/journal.pgen.1005166](https://doi.org/10.1371/journal.pgen.1005166)
- [45] Ogorodnikov A, Levin M, Tattikota S, et al. Transcriptome 3' end organization by PCF11 links alternative polyadenylation to formation and neuronal differentiation of neuroblastoma. *Nat Commun*. 2018;9(1). doi: [10.1038/s41467-018-07580-5](https://doi.org/10.1038/s41467-018-07580-5)
- [46] Turner RE, Henneken LM, Liem-Weits M, et al. Requirement for cleavage factor IIm in the control of alternative polyadenylation in breast cancer cells. *RNA*. 2020;26(8):969–981. doi: [10.1261/rna.075226.120](https://doi.org/10.1261/rna.075226.120)
- [47] Zhu Y, Wang X, Forouzmand E, et al. Molecular mechanisms for CFIm-mediated regulation of mRNA alternative polyadenylation. *Mol Cell*. 2018;69(1):62–74.e4. doi: [10.1016/j.molcel.2017.11.031](https://doi.org/10.1016/j.molcel.2017.11.031)
- [48] Xia Z, Donehower LA, Cooper TA, et al. Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types. *Nat Commun*. 2014;5(1):5274. doi: [10.1038/ncomms6274](https://doi.org/10.1038/ncomms6274)
- [49] Davis AG, Johnson DT, Zheng D, et al. Alternative polyadenylation dysregulation contributes to the differentiation block of acute myeloid leukemia. *Blood*. 2022;139(3):424–438. doi: [10.1182/blood.2020005693](https://doi.org/10.1182/blood.2020005693)
- [50] Caggiano C, Pieraccioli M, Pitolli C, et al. The androgen receptor couples promoter recruitment of RNA processing factors to regulation of alternative polyadenylation at the 3' end of transcripts. *Nucleic Acids Res*. 2022;50(17):9780–9796. doi: [10.1093/nar/gkac737](https://doi.org/10.1093/nar/gkac737)
- [51] Li W, Li W, Laishram RS, et al. Distinct regulation of alternative polyadenylation and gene expression by nuclear poly(A) polymerases. *Nucleic Acids Res*. 2017;45(15):8930–8942. doi: [10.1093/nar/gkx560](https://doi.org/10.1093/nar/gkx560)
- [52] Gruber AR, Martin G, Keller W, et al. Cleavage factor im is a key regulator of 3' UTR length. *RNA Biol*. 2012;9(12):1405–1412. doi: [10.4161/rna.22570](https://doi.org/10.4161/rna.22570)
- [53] Richard P, Vethantham V, Manley JL. Roles of sumoylation in mRNA processing and metabolism BT - SUMO regulation of cellular processes. V. G. Wilson, edited by. Cham: Springer International Publishing; 2017. pp. 15–33.
- [54] Ji Z, Luo W, Li W, et al. Transcriptional activity regulates alternative cleavage and polyadenylation. *Mol Syst Biol*. 2011;7(1):534. doi: [10.1038/msb.2011.69](https://doi.org/10.1038/msb.2011.69)
- [55] Liu X, Freitas J, Zheng D, et al. Transcription elongation rate has a tissue-specific impact on alternative cleavage and polyadenylation in *Drosophila melanogaster*. *RNA*. 2017;23(12):1807–1816. doi: [10.1261/rna.062661.117](https://doi.org/10.1261/rna.062661.117)
- [56] Pinto PAB, Henriques T, Freitas MO, et al. RNA polymerase II kinetics in polo polyadenylation signal selection. *EMBO J*. 2011;30(12):2431–2444. doi: [10.1038/emboj.2011.156](https://doi.org/10.1038/emboj.2011.156)
- [57] LaBella ML, Hujber EJ, Moore KA, et al. Casein kinase 1δ stabilizes mature axons by inhibiting transcription termination of Ankyrin. *Dev Cell*. 2020;52(1):88–103.e18. doi: [10.1016/j.devcel.2019.12.005](https://doi.org/10.1016/j.devcel.2019.12.005)
- [58] Alfonso-Gonzalez C, Legnini I, Holec S, et al. Sites of transcription initiation drive mRNA isoform selection. *Cell*. 2023;186(11):2438–2455.e22. doi: [10.1016/j.cell.2023.04.012](https://doi.org/10.1016/j.cell.2023.04.012)
- [59] Nimura K, Yamamoto M, Takeichi M, et al. Regulation of alternative polyadenylation by Nkx2-5 and Xrn2 during mouse heart development. *Elife*. 2016;5:e16030.
- [60] You L-Y, Lin J, Xu H-W, et al. Intragenic heterochromatin-mediated alternative polyadenylation modulates miRNA and pollen development in rice. *New Phytol*. 2021;232(2):835–852. doi: [10.1111/nph.17635](https://doi.org/10.1111/nph.17635)
- [61] Kwon B, Fansler MM, Patel ND, et al. Enhancers regulate 3' end processing activity to control expression of alternative 3'UTR isoforms. *Nat Commun*. 2022;13(1):2709. doi: [10.1038/s41467-022-30525-y](https://doi.org/10.1038/s41467-022-30525-y)
- [62] Mayr C, Bartel DP. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell*. 2009;138(4):673–684. doi: [10.1016/j.cell.2009.06.016](https://doi.org/10.1016/j.cell.2009.06.016)
- [63] Mansfield KD, Keene JD. Neuron-specific ELAV/Hu proteins suppress HuR mRNA during neuronal differentiation by alternative polyadenylation. *Nucleic Acids Res*. 2012;40(6):2734–2746. doi: [10.1093/nar/gkr1114](https://doi.org/10.1093/nar/gkr1114)
- [64] Ciolli Mattioli C, Rom A, Franke V, et al. Alternative 3' UTRs direct localization of functionally diverse protein isoforms in neuronal compartments. *Nucleic Acids Res*. 2019;47(5):2560–2573. doi: [10.1093/nar/gky1270](https://doi.org/10.1093/nar/gky1270)
- [65] Sandberg R, Neilson JR, Sarma A, et al. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer MicroRNA target sites. *Science*. 2008;80-. 320(5883):1643–1647. doi: [10.1126/science.1155390](https://doi.org/10.1126/science.1155390)
- [66] Domingues RG, Lago-Baldaia I, Pereira-Castro I, et al. CD5 expression is regulated during human T-cell activation by alternative polyadenylation, PTBP1, and miR-204. *Eur J Immunol*. 2016;46(6):1490–1503. doi: [10.1002/eji.201545663](https://doi.org/10.1002/eji.201545663)
- [67] Gruber AR, Martin G, Müller P, et al. Global 3' UTR shortening has a limited effect on protein abundance in proliferating T cells. *Nat Commun*. 2014;5(1):5465. doi: [10.1038/ncomms6465](https://doi.org/10.1038/ncomms6465)
- [68] Li W, Park JY, Zheng D, et al. Alternative cleavage and polyadenylation in spermatogenesis connects chromatin regulation with post-transcriptional control. *BMC Biol*. 2016;14(1). doi: [10.1186/s12915-016-0229-6](https://doi.org/10.1186/s12915-016-0229-6)
- [69] Shan L, Wu C, Chen D, et al. Regulators of alternative polyadenylation operate at the transition from mitosis to meiosis. *J Genet Genomics*. 2017;44(2):95–106. doi: [10.1016/j.jgg.2016.12.007](https://doi.org/10.1016/j.jgg.2016.12.007)
- [70] Lau AG, Irier HA, Gu J, et al. Distinct 3'UTRs differentially regulate activity-dependent translation of brain-derived neurotrophic factor (BDNF). *Proc Natl Acad Sci*. 2010;107(36):15945–15950. doi: [10.1073/pnas.1002929107](https://doi.org/10.1073/pnas.1002929107)
- [71] Blair JD, Hockemeyer D, Doudna JA, et al. Widespread Translational Remodeling during Human Neuronal Differentiation. *Cell Rep*. 2017;21(7):2005–2016. doi: [10.1016/j.celrep.2017.10.095](https://doi.org/10.1016/j.celrep.2017.10.095)
- [72] Berkovits BD, Mayr C. Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. *Nature*. 2015;522(7556):363–367. doi: [10.1038/nature14321](https://doi.org/10.1038/nature14321)
- [73] Pelka GJ, Watson CM, Christodoulou J, et al. Distinct expression profiles of Mecp2 transcripts with different lengths of 3'UTR in the brain and visceral organs during mouse development. *Genomics*. 2005;85(4):441–452. doi: [10.1016/j.ygeno.2004.12.002](https://doi.org/10.1016/j.ygeno.2004.12.002)
- [74] Mullaney BC, Johnston MV, Blue ME. Developmental expression of methyl-CpG binding protein 2 is dynamically regulated in the rodent brain. *Neuroscience*. 2004;123(4):939–949. doi: [10.1016/j.neuroscience.2003.11.025](https://doi.org/10.1016/j.neuroscience.2003.11.025)
- [75] Costessi L, Devescovi G, Baralle FE, et al. Brain-specific promoter and polyadenylation sites of the -adducin pre-mRNA generate an unusually long 3'-UTR. *Nucleic Acids Res*. 2006;34(1):243–253. doi: [10.1093/nar/gkj425](https://doi.org/10.1093/nar/gkj425)
- [76] Hilgers V, Perry MW, Hendrix D, et al. Neural-specific elongation of 3' UTRs during *Drosophila* development. *Proc Natl Acad Sci*. 2011;108(38):15864–15869. doi: [10.1073/pnas.1112672108](https://doi.org/10.1073/pnas.1112672108)
- [77] Sanfilippo P, Wen J, Lai EC. Landscape and evolution of tissue-specific alternative polyadenylation across *Drosophila* species. *Genome Biol*. 2017;18(1):229. doi: [10.1186/s13059-017-1358-0](https://doi.org/10.1186/s13059-017-1358-0)

- [78] Smibert P, Miura P, Westholm JO, et al. Global patterns of tissue-specific alternative polyadenylation in *Drosophila*. *Cell Rep.* 2012;1(3):277–289. doi: [10.1016/j.celrep.2012.01.001](#)
- [79] Ji Z, Lee JY, Pan Z, et al. Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. *Proc Natl Acad Sci.* 2009;106(17):7028–7033. doi: [10.1073/pnas.0900028106](#)
- [80] Shepard PJ, Choi EA, Lu J, et al. Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. *RNA.* 2011;17(4):761–772. doi: [10.1261/rna.2581711](#)
- [81] Ulitsky I, Shkumatava A, Jan CH, et al. Extensive alternative polyadenylation during zebrafish development. *Genome Res.* 2012;22(10):2054–2066. doi: [10.1101/gr.139733.112](#)
- [82] Tushev G, Glock C, Heumüller M, et al. Alternative 3' UTRs modify the localization, regulatory potential, stability, and plasticity of mRNAs in neuronal compartments. *Neuron.* 2018;98(3):495–511.e6. doi: [10.1016/j.neuron.2018.03.030](#)
- [83] Vicario A, Colliva A, Ratti A, et al. Dendritic targeting of short and long 3' UTR BDNF mRNA is regulated by BDNF or NT-3 and distinct sets of RNA-binding proteins. *Front Mol Neurosci.* 2015;8. doi: [10.3389/fnmol.2015.00062](#)
- [84] Yudin D, Hanz S, Yoo S, et al. Localized regulation of axonal RanGTPase controls retrograde injury signaling in peripheral nerve. *Neuron.* 2008;59(2):241–252. doi: [10.1016/j.neuron.2008.05.029](#)
- [85] Bae B, Gruner HN, Lynch M, et al. Elimination of *Calm1* long 3'-UTR mRNA isoform by CRISPR-Cas9 gene editing impairs dorsal root ganglion development and hippocampal neuron activation in mice. *RNA.* 2020;26(10):1414–1430. doi: [10.1261/rna.076430.120](#)
- [86] Andreassi C, Luisier R, Crerar H, et al. Cytoplasmic cleavage of *IMPA1* 3' UTR is necessary for maintaining axon integrity. *Cell Rep.* 2021;34(8):108778. doi: [10.1016/j.celrep.2021.108778](#)
- [87] Crerar H, Scott-Solomon E, Bodkin-Clarke C, et al. Regulation of NGF Signaling by an Axonal Untranslated mRNA. *Neuron.* 2019;102(3):553–563.e8. doi: [10.1016/j.neuron.2019.02.011](#)
- [88] Li H, Janssens J, De Waegeneer M, et al. Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly. *Science.* 2023a;80-.). 375:eabk2432.
- [89] Lianoglou S, Garg V, Yang JL, et al. Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes Dev.* 2013;27(21):2380–2396. doi: [10.1101/gad.229328.113](#)
- [90] Hoque M, Ji Z, Zheng D, et al. Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. *Nat Methods.* 2013;10(2):133–139. doi: [10.1038/nmeth.2288](#)
- [91] Ji Z, Tian B, Valcarcel J. Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. *PLoS One.* 2009;4(12):e8419. doi: [10.1371/journal.pone.0008419](#)
- [92] Vest KE, Paskavitz AL, Lee JB, et al. Dynamic changes in copper homeostasis and post-transcriptional regulation of *Atp7a* during myogenic differentiation†. *Metallomics.* 2018;10(2):309–322. doi: [10.1039/C7MT00324B](#)
- [93] Raz V, Riaz M, Tatum Z, et al. 't Hoen, 2018 the distinct transcriptomes of slow and fast adult muscles are delineated by non-coding RNAs. *FASEB J.* 2018;32(3):1579–1590. doi: [10.1096/fj.201700861R](#)
- [94] Kuang S, Kuroda K, Le Grand F, et al. Asymmetric self-renewal and commitment of satellite stem cells in muscle. *Cell.* 2007;129(5):999–1010. doi: [10.1016/j.cell.2007.03.044](#)
- [95] Buckingham M, Relaix F. PAX3 and PAX7 as upstream regulators of myogenesis. *Semin Cell Dev Biol.* 2015;44:115–125. doi: [10.1016/j.semdb.2015.09.017](#)
- [96] Boutet SC, Cheung TH, Quach NL, et al. Alternative polyadenylation mediates MicroRNA regulation of muscle stem cell function. *Cell Stem Cell.* 2012;10(3):327–336. doi: [10.1016/j.stem.2012.01.017](#)
- [97] Vorobyov E, Horst J. Expression of two protein isoforms of PAX7 is controlled by competing cleavage-polyadenylation and splicing. *Gene.* 2004;342(1):107–112. doi: [10.1016/j.gene.2004.07.030](#)
- [98] de Morree A, Klein JDD, Gan Q, et al. Alternative polyadenylation of Pax3 controls muscle stem cell fate and muscle function. *Science.* 2019;80-.). 366(6466):734–738. doi: [10.1126/science.aax1694](#)
- [99] Anvar SY, Raz Y, Verwey N, et al. A decline in PABPN1 induces progressive muscle weakness in Oculopharyngeal muscle dystrophy and in muscle aging. *Aging.* 2013;5(6):412–426. doi: [10.18632/aging.100567](#)
- [100] Mei H, Boom J, el Abdellaoui S, et al. Alternative polyadenylation utilization results in Ribosome assembly and mRNA translation deficiencies in a model for muscle aging. *J Gerontol Ser A.* 2022;77(6):1130–1140. doi: [10.1093/gerona/glac058](#)
- [101] Smith SS, Kessler CB, Shenoy V, et al. IGF-I 3' untranslated region: strain-specific polymorphisms and motifs regulating IGF-I in Osteoblasts. *Endocrinology.* 2013;154(1):253–262. doi: [10.1210/en.2012-1476](#)
- [102] Khajuria DK, Nowak I, Leung M, et al. Transcript shortening via alternative polyadenylation promotes gene expression during fracture healing. *Bone Res.* 2023;11(1):5. doi: [10.1038/s41413-022-00236-7](#)
- [103] Edwalds-Gilbert G, Veraldi KL, Milcarek C. Alternative poly(A) site selection in complex transcription units: means to an end? *Nucleic Acids Res.* 1997;25(13):2547–2561. doi: [10.1093/nar/25.13.2547](#)
- [104] Peterson ML. Immunoglobulin heavy chain gene regulation through polyadenylation and splicing competition. *WIREs RNA.* 2011;2(1):92–105. doi: [10.1002/wrna.36](#)
- [105] Rogers J, Early P, Carter C, et al. Two mRNAs with different 3' ends encode membrane-bound and secreted forms of immunoglobulin μ chain. *Cell.* 1980;20(2):303–312. doi: [10.1016/0092-8674\(80\)90616-9](#)
- [106] Takagaki Y, Seipelt RL, Peterson ML, et al. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell.* 1996;87(5):941–952. doi: [10.1016/S0092-8674\(00\)82000-0](#)
- [107] Takagaki Y, Manley JL. Levels of polyadenylation factor CstF-64 control IgM heavy chain mRNA accumulation and other events associated with B cell differentiation. *Mol Cell.* 1998;2(6):761–771. doi: [10.1016/S1097-2765\(00\)80291-9](#)
- [108] Edwalds-Gilbert G, Milcarek C. Regulation of poly(A) site use during mouse B-Cell development involves a change in the binding of a general polyadenylation factor in a B-Cell stage-specific manner. *Mol Cell Biol.* 1995;15(11):6420–6429. doi: [10.1128/MCB.15.11.6420](#)
- [109] Martincic K, Campbell R, Edwalds-Gilbert G, et al. Increase in the 64-kDa subunit of the polyadenylation/cleavage stimulatory factor during the G0 to S phase transition. *Proc Natl Acad Sci.* 1998;95(19):11095–11100. doi: [10.1073/pnas.95.19.11095](#)
- [110] Pai AA, Baharian G, Pagé Sabourin A, et al. Widespread shortening of 3' untranslated regions and increased exon inclusion are evolutionarily conserved features of innate immune responses to infection. *PLoS Genet.* 2016;12(9):e1006338. doi: [10.1371/journal.pgen.1006338](#)
- [111] Jia X, Yuan S, Wang Y, et al. The role of alternative polyadenylation in the antiviral innate immune response. *Nat Commun.* 2017;8(1):14605. doi: [10.1038/ncomms14605](#)
- [112] Kalam H, Fontana MF, Kumar D, et al. Alternate splicing of transcripts shape macrophage response to mycobacterium tuberculosis infection. *PLoS Pathog.* 2017;13(3):e1006236. doi: [10.1371/journal.ppat.1006236](#)
- [113] Chuvpilo S, Zimmer M, Kerstan A, et al. Alternative polyadenylation events contribute to the induction of NF-ATc in effector T cells. *Immunity.* 1999;10(2):261–269. doi: [10.1016/S1074-7613\(00\)80026-6](#)

- [114] Li K, Qiang M, Xu Y. Cell-type-specific alternative polyadenylation as a therapeutic biomarker in lung cancer progression. *Mol Ther Nucleic Acids*. 2023b;34:102030. doi: [10.1016/j.omtn.2023.102030](https://doi.org/10.1016/j.omtn.2023.102030)
- [115] Venkat S, Tisdale AA, Schwarz JR, et al. Alternative polyadenylation drives oncogenic gene expression in pancreatic ductal adenocarcinoma. *Genome Res*. 2020;30(3):347–360. doi: [10.1101/gr.257550.119](https://doi.org/10.1101/gr.257550.119)
- [116] Ye W, Lian Q, Ye C, et al. A survey on Methods for predicting polyadenylation sites from DNA sequences, bulk RNA-seq, and single-cell RNA-seq. *Genomics Proteomics Bioinf*. 2023;21(1):67–83. doi: [10.1016/j.gpb.2022.09.005](https://doi.org/10.1016/j.gpb.2022.09.005)
- [117] Legnini I, Alles J, Karaïskos N, et al. FLAM-seq: full-length mRNA sequencing reveals principles of poly(A) tail length control. *Nat Methods*. 2019;16(9):879–886. doi: [10.1038/s41592-019-0503-y](https://doi.org/10.1038/s41592-019-0503-y)
- [118] Wang J, Chen W, Yue W, et al. Comprehensive mapping of alternative polyadenylation site usage and its dynamics at single-cell resolution. *Proc Natl Acad Sci, USA*. 2022;119(49):e2113504119. doi: [10.1073/pnas.2113504119](https://doi.org/10.1073/pnas.2113504119)
- [119] Kim N, Chung W, Eum HH, et al. Alternative polyadenylation of single cells delineates cell types and serves as a prognostic marker in early stage breast cancer. *PLoS One*. 2019;14(5):e0217196. doi: [10.1371/journal.pone.0217196](https://doi.org/10.1371/journal.pone.0217196)
- [120] Ye C, Zhou Q, Hong Y, et al. Role of alternative polyadenylation dynamics in acute myeloid leukaemia at single-cell resolution. *RNA Biol*. 2019;16(6):785–797. doi: [10.1080/15476286.2019.1586139](https://doi.org/10.1080/15476286.2019.1586139)