

Processing and transcriptome expansion at the mRNA 3' end in health and disease: finding the right end

Anton Ogorodnikov^{1,2} · Yulia Kargapolova^{1,2} · Sven Danckwardt^{1,2,3}

Received: 12 April 2016 / Accepted: 19 April 2016 / Published online: 25 May 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract The human transcriptome is highly dynamic, with each cell type, tissue, and organ system expressing an ensemble of transcript isoforms that give rise to considerable diversity. Apart from alternative splicing affecting the “body” of the transcripts, extensive transcriptome diversification occurs at the 3' end. Transcripts differing at the 3' end can have profound physiological effects by encoding proteins with distinct functions or regulatory properties or by affecting the mRNA fate via the inclusion or exclusion of regulatory elements (such as miRNA or protein binding sites). Importantly, the dynamic regulation at the 3' end is associated with various (patho) physiological processes, including the immune regulation but also tumorigenesis. Here, we recapitulate the mechanisms of constitutive mRNA 3' end processing and review the current understanding of the dynamically regulated diversity at the transcriptome 3' end. We illustrate the medical importance by presenting examples that are associated with perturbations of this process and indicate resulting implications for molecular diagnostics as well as potentially arising novel therapeutic strategies.

Keywords Alternative cleavage and polyadenylation (APA) · Disease · Post-transcriptional gene regulation · RNA 3' end maturation · Transcriptome diversity · Cancer

✉ Sven Danckwardt
Sven.Danckwardt@unimedizin-mainz.de

¹ Institute for Clinical Chemistry and Laboratory Medicine, University Medical Center Mainz, Langenbeckstr 1, 55131 Mainz, Germany

² Center for Thrombosis and Hemostasis (CTH), University Medical Center Mainz, Langenbeckstr 1, 55131 Mainz, Germany

³ German Center for Cardiovascular Research (DZHK), Langenbeckstr 1, 55131 Mainz, Germany

Introduction

The mRNA and protein isoforms produced by alternative processing of primary RNA transcripts differ in structure, function, localization, or other properties [13, 120, 183]. Alternative splicing affects more than half of all human genes and represents a primary driver of the evolution of phenotypic complexity in mammals [15, 80, 95]. Across individuals, changes in normal isoform structure can have phenotypic consequences and have been associated with disease [52]. With the advent of next-generation RNA sequencing technologies, it became apparent that not only the “body” of transcripts but also the mRNA 3' end is affected by enormous diversity. Up to 70 % of the transcriptome undergoes alternative mRNA 3' end processing [92]. This for example affects numerous genes during the stress response [187] or after T and B cell activation [152]. The medical significance is highlighted by complex disorders that are associated with alternative cleavage and polyadenylation (APA), i.e., in the susceptibility to systemic lupus erythematosus [62] or more globally in tumorigenesis [119, 121]. Yet, up to now, the functional role of widespread APA in disease processes is still enigmatic.

In the following, we will briefly present the mechanistic key features of 3' end formation since several reviews cover the basics of 3' end formation in greater detail [28, 43, 60, 88, 143, 180, 185, 198]. We will illustrate the medical perspective of 3' end processing and show how disease can be caused by perturbations of canonical mRNA processing in *cis* and *trans*. We will then present how alternative 3' end processing contributes to the complexity of the transcriptome and thereby affects important cellular functions in physiological as well as pathophysiological conditions. Finally and most intriguingly, we will discuss to what extent alternative cleavage and polyadenylation represent the driver or passenger in the pathogenesis of human disorders.

General principles of canonical 3' end processing—the eukaryotic mRNA 3' end cleavage and polyadenylation machinery

All eukaryotic mRNAs, except some replication dependent histone mRNAs, as well as several non-coding transcripts including miRNAs, possess poly(A)-tails at their 3' end, which are produced by a relatively simple two-step reaction involving endonucleolytic cleavage and subsequent non-templated poly(A) tail addition (Fig. 1; reviewed [127, 142, 156, 192]). The executing molecular machinery however is complex. It involves more than 50 proteins [160], which are loaded onto more or less highly conserved sequence motifs to catalyze this step of pre-mRNA maturation. The specificity and efficiency of this process are determined by the binding of two core multi-protein complexes (CPSF and CSTF, for “cleavage and polyadenylation specificity factor” and “cleavage stimulation factor”) to sequences surrounding the poly(A) site as soon as the nascent pre-mRNA transcript emerges from the elongating RNA polymerase II (POL2). After assembling the 3' end processing apparatus, the cleavage reaction is catalyzed by CPSF-73 [45, 112, 149] and the cleaved mRNA is polyadenylated by the nuclear poly(A)-polymerase (PAP) adding ~50–100 A-nucleotides to the 3' end [24]. The length of the poly(A) tail is similar in different mRNAs and is determined by an interaction between the nuclear poly(A) binding protein (PABPN1), CPSF, and PAP [91]. Upon nuclear export, PABPN1 is replaced by the cytosolic poly(A) binding protein (PABPC), which establishes contact with the translation initiation factor eIF4G, stimulating translation and regulating mRNA stability [82, 91, 150]. Thus, 3' end polyadenylation is vital for various steps of gene regulation; it is required for nuclear export and stability of mature transcripts and for efficient translation of mRNAs [151]. Interestingly, although key aspects of mRNA 3' end maturation have been identified decades ago, recent studies have significantly reshaped current models for the protein-RNA interactions involved in the poly(A) site (PAS) recognition [161]. This, includes the recognition of the highly conserved AAUAAA hexanucleotide motif (and less commonly AUUAAA) defining the site of 3' end processing of almost all polyadenylated mRNAs [23, 154] (see below).

mRNA 3' end processing is tightly connected to transcription and splicing to ensure proper mRNA maturation and to maintain genome integrity

Messenger RNA 3' end processing is well-orchestrated and interconnected with transcription, splicing, and translation [70, 114, 143] (Fig. 1b): As the pre-mRNA transcript emerges

from POL2, extensive constitutive and alternative splicing events occur co-transcriptionally giving rise to a perplexingly high transcriptome diversity. Eventually, transcription termination pauses the elongating polymerase triggered by recognition of poly(A) signals in the nascent transcript [143, 148], and both CPSF and CSTF are transferred by the carboxy-terminal domain (CTD) of POL2 to their specific pre-mRNA binding sites to produce the mRNA 3' end. The phosphorylation of serine and threonine residues within the CTD regulates gene expression [1]; it coordinates the recruitment of RNA processing factors (including the 5' capping complex, the spliceosome, and 3' processing machinery) and regulates chromatin organization by histone methylation [68].

The extensive crosstalk between the different pre-mRNA processing activities (capping, splicing, and polyadenylation) is crucial for gene expression and, importantly, genome integrity (Fig. 1c): Proteins binding to the cap structure of pre-mRNAs interact with splicing factors and promote recognition of the cap-proximal splice site. Conversely, splicing factors associating with the 3' terminal intron interact with downstream polyadenylation factors to mutually promote 3' end cleavage/polyadenylation and terminal intron splicing [5, 36, 65, 93, 100, 109, 122, 125, 126, 134, 175, 184]. These interactions ensure the recognition of the correct splice sites and the timely, accurate, and efficient 3' end processing. Moreover, the extensive integration between the different co-transcriptional mechanisms protects chromosomes from potentially deleterious effects, which could arise from interaction between the nascent RNA and template DNA during transcription [101]. Finally, functional polyadenylation sites and polyadenylation factors are required for efficient transcription termination [20, 85, 143, 148] and release of the polyadenylated mRNAs for export from the nucleus [145]. The efficiency of polyadenylation can thus have significant quantitative effects on gene expression in general, and defects of mRNA 3' end formation can profoundly affect cell viability, growth, and development.

The medical relevance of errors of 3' end processing is exemplified by different inherited and acquired human disorders. A continuously growing number of reports document the increasing awareness of the mRNA 3' end becoming a critical constituent for a variety of disorders (for review [34]).

In the following paragraphs, we highlight the most characteristic hallmarks of constitutive 3' end processing and—based on a few selected examples—illustrate how alterations of sequence elements or the executing 3' end processing machinery can result in human pathology. In the second part, we shift to regulated and alternative 3' end formation and demonstrate its importance in various physiologically and pathophysiologically relevant processes.

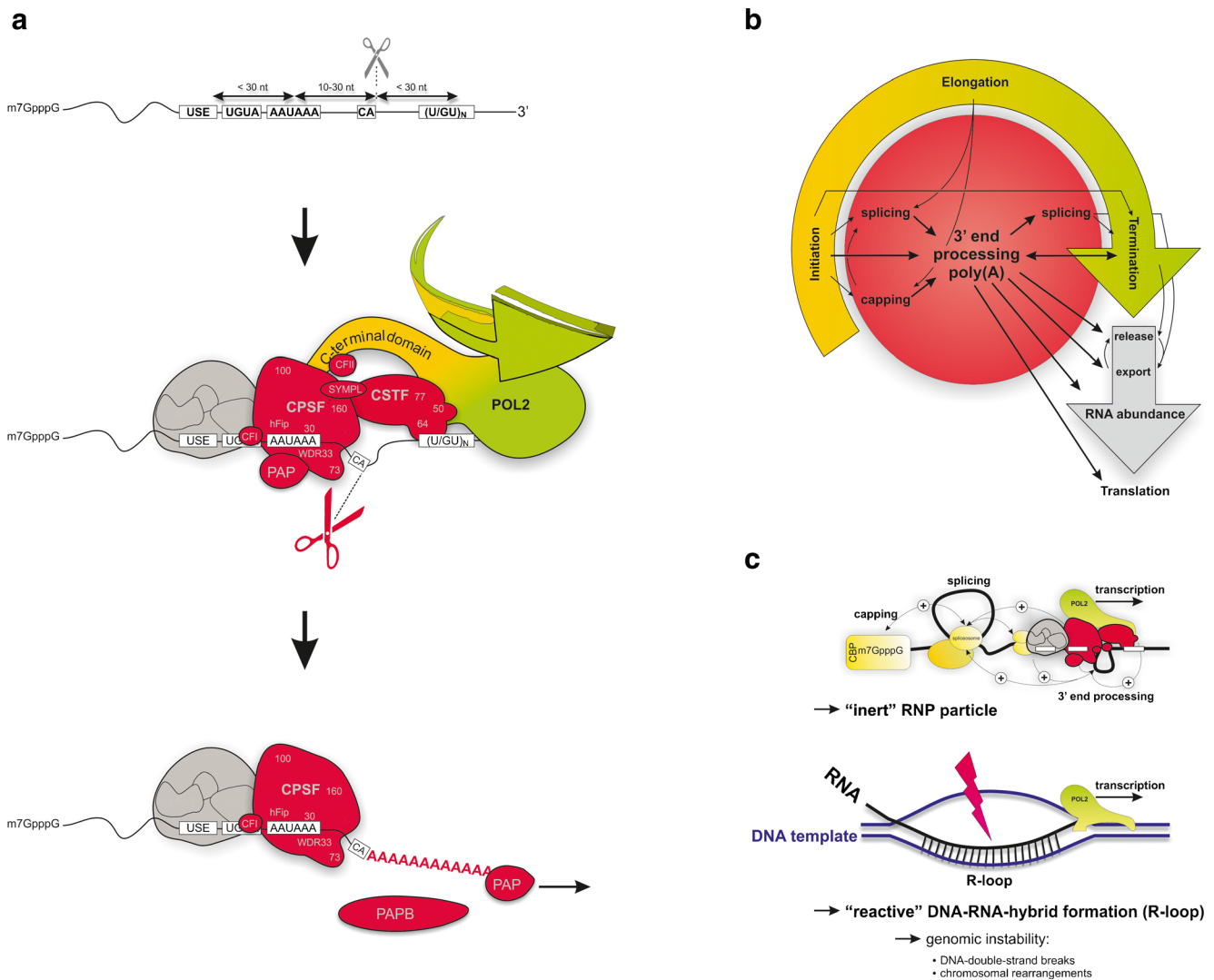


Fig. 1 Sequence elements and protein components involved in the formation of poly(A) tails. **a** The specificity and efficiency of cleavage and polyadenylation are determined by the binding of multi-protein complexes to specific elements at the 3' end of the pre-mRNA. Most pre-mRNAs contain two core elements. The canonical polyadenylation signal AAUAAA (or less frequently AUUAAA) upstream of the cleavage site is recognized by the multimeric cleavage and polyadenylation specificity factor (CPSF) consisting of at least six subunits (CPSF 160, CPSF 100, CPSF 73, CPSF 30, hFlp1, and WDR33). This RNA-protein interaction determines the site of cleavage 10–30 nt downstream, preferentially immediately 3' of a CA dinucleotide. The second canonical sequence element is characterized by a high density of G/U or U residues which is located up to 30 nt downstream of the cleavage site. This downstream sequence element (DSE) is bound by the 64-kDa subunit of the hetero-trimeric cleavage stimulating factor (CSTF) that promotes the efficiency of 3' end processing. Furthermore, multimers of a UGUA motif are localized at variable distances upstream of the cleavage site to recruit the heterodimeric cleavage factor CFI [19, 176]. Finally, accessory sequences can function as upstream sequence elements (USE) [17, 32, 36, 66, 67, 75, 98, 128, 129, 132, 193] to facilitate 3' end processing by recruiting canonical 3' end factors directly [128, 129] or by serving as an additional anchor for the (canonical) 3' end processing machinery [36, 66]. After assembly of the

basal 3' end processing machinery, the endonucleolytic cleavage reaction is catalyzed by CPSF-73, and the cleaved mRNA is polyadenylated by nuclear poly(A)-polymerase (PAP). The binding of PABPN1 to the poly(A) tail is unstable, and upon nuclear export, PABPN1 is replaced by the cytosolic poly(A) binding protein (PABPC), which interacts with the translation initiation factor eIF4G, stimulating translation and regulating mRNA stability. Of note, in addition to histone 3' end processing that follows a different pathway [118, 185], other “non-canonical” mechanisms of 3' end processing exist [189]. **b** Transcription initiation, elongation, and termination (*circular arrow*) are tightly coupled to mRNA processing steps such as capping, splicing, and 3' end processing (*inner circle*). Appropriate 3' end processing is functionally interconnected with transcription and mRNA capping and splicing and impacts on post-transcriptional mechanisms (mRNA release, export, abundance, and translation). Loss- or gain-of-function of 3' end processing thus critically interferes with other gene expression steps (modified after [34]). **c** Co-transcriptional mRNA processing promotes packaging of the nascent RNA transcript (formation of an “inert” RNP particle) and thus to prevent the accumulation of co-transcriptional R-loops (*lower panel*), which can lead to DNA-double-strand breaks and chromosomal rearrangements. Disruption of co-transcriptional RNA-processing can thus lead to genomic instability (modified after [34])

When processing gets awry

The polyadenylation signal matters

Numerous mutations affecting *cis*-regulatory sequence elements required for mRNA 3' end processing evidence their detrimental role in a variety of human disorders [34]. Intriguingly, a significant factor in deciphering the underlying mechanistic principles of mRNA 3' end processing was the presence of mutations, which were initially identified in thalassemia patients. Among those patients, different mutations affecting the AAUAAA hexanucleotide of the poly(A) signal were found in the α - [69] and β -globin genes [136] (and references therein) and shown to invariably inactivate or severely inhibit gene expression. Similar mutations were observed in, e.g., the *Foxp3* gene causing IPEX syndrome [10], a rare fatal disorder characterized by immune dysregulation, polyendocrinopathy, enteropathy, and X-linked inheritance. Recently, a germline variant in the TP53 polyadenylation signal has been identified in a large genome-wide association study to confer cancer susceptibility. In this gene, a mutation that changes the AATAAA into AATACA results in impaired 3' end processing of the TP53 mRNA predisposing to prostate cancer, gliomas, and colorectal adenomas [168].

These and other examples [34] illustrate the functional importance of the highly conserved poly(A) signal and reveal the devastating consequences of some of these mutations. Although there is some sequence flexibility [9, 173] and even alternative 3' UTR architectures for effective processing [135], these findings indicate that the hexanucleotide required for recruitment of the CPSF complex represents the “Achilles heel” for disease causing loss-of-function mutations altering 3' end processing. This is further corroborated by numerous mutations altering polyadenylation in a variety of other disorders (for review [34]).

How about other sequence elements?

3' end processing depends on various canonical and non-canonical (auxiliary) sequence elements (Fig. 1a). With the exception of the poly(A) hexanucleotide, they are less well conserved and consequently tolerate (“silent”) nucleotide exchanges with greater likelihood.

Nevertheless, examples, which illustrate the function of these elements, exist. This has been first demonstrated for a clinically relevant gain of function mutation of the cleavage site stimulating 3' end processing [59]: In most cases, endonucleolytic cleavage and polyadenylation of pre-mRNAs occurs predominantly 3' of a CA dinucleotide. However, in the prothrombin (*F2*) gene encoding a key blood coagulation factor, the cleavage site is composed of a CG dinucleotide, which is less efficient in promoting the cleavage reaction [26]. As a consequence, a common mutation (*F2* 20210*A) affecting the

most 3' nucleotide of the *F2* mRNA converts the physiologically inefficient cleavage site into the mechanistically most efficient CA dinucleotide [32, 59]. This increases the cleavage site recognition resulting in an approximately twofold enhancement of *F2* mRNA and protein expression. This finally causes raised *F2* plasma concentrations, which disturb the finely tuned balance between pro- and anti-coagulatory activities and thereby predisposes to thrombosis [141].

Mutations affecting 3' end processing were also identified at other positions in the *F2* gene [6, 155, 165]. They increase the efficiency of 3' end processing [33] either when located at the penultimate position of the *F2* 3'UTR (*F2* 20209*T) or further downstream 3' of the cleavage site in the putative CSTF binding site (*F2* 20221*T). However, these effects are presumably gene-specific since the putative CSTF binding site in the *F2* gene displays an unusually low density of uridine residues when compared to efficiently 3' end processed mRNAs. Consequently, the introduction of (an) additional uridine-residue(s) into that region enhances 3' end processing, supposedly by facilitating the interaction of CSTF with the pre-mRNA [32]. This illustrates a typical feature of most PASs that typically harbor a U-/GU-rich stretch up to 30 nucleotides downstream of the cleavage site to recruit the CSTF complex for efficient 3' end cleavage and polyadenylation (Fig. 1a).

Yet deleterious effects arising from mutations have been documented, which neither directly affect the poly(A) signal or the cleavage site nor the downstream sequence elements. This, for instance, is shown for a 20 base pair duplication in *MSH6*, one of the four mismatch repair genes causing the *LYNCH* syndrome (HNPCC or hereditary nonpolyposis colorectal cancer), an autosomal-dominant genetic cancer syndrome with a high risk of colon cancer (among others) [41]. This duplication downregulates processing at an adjacent PAS, although the underlying mechanism remained unclear. Another intriguing example highlighting the mechanistic complexity of mRNA 3' end processing is found for a complex immunodeficiency syndrome, which can be caused by a 3'UTR mutation in the *p14* gene. In this case, the mutation creates a splicing defective 5' splice site resulting in a U1 snRNP-mediated suppression of an adjacent PAS [96]. Interestingly, this principle already points to a mechanistic aspect that will be relevant in the regulation of alternative cleavage and polyadenylation and the therapeutic manipulation of 3' end processing (see below).

These examples suggest that aberrations in regions not directly affecting the poly(A) signal deserve special attention. Although there clearly is more sequence flexibility up- and downstream of the poly(A) signal, mutations in such regions can have devastating consequences. This aspect points to another mechanistic peculiarity of 3' end processing that the majority of RNAs likely do not have “optimal” upstream and downstream core elements. Instead, auxiliary elements

situated in this regions aid polyadenylation (for review, [37]). However, their composition and mode of action are often gene-specific thereby accommodating the needs for specificity of regulation at the mRNA 3' end (further detailed below).

Role of *trans*-acting factors

As indicated before, the processing apparatus is a complex machinery involving more than 50 proteins [160]. Examples illustrating the medical relevance of mutations affecting 3' end processing factors are highlighted by oculopharyngeal muscular dystrophy (OPMD) or the hypereosinophilic syndrome (HES).

OPMD is an adult-onset disease with slowly progressing muscle weakness primarily affecting the eyelids resulting in ptosis and the pharyngeal muscles resulting in dysphagia. It is caused by short trinucleotide repeat [(GCG)_{8–13}] expansions in the coding region of the nuclear poly(A)binding protein 1 (PABPN1, see above) [18]. Normally, the polyalanine stretch encoded by this trinucleotide comprises 10 alanines, which is expanded to 12–17 alanines in autosomal-dominant OPMD. This expansion results in an increase of self-association, misfolding, and filamentous nuclear aggregation of the PABPN1 protein in skeletal muscle. In vitro, the mutant protein is fully active and OPMD cells do not display a severe polyadenylation defect [21, 91]. Thus, the phenotype might be best explained by either a quantitatively minor disturbance of the protein's function in polyadenylation (which may be difficult to detect in vitro or in transfected cells) and/or by co-sequestration of other potentially interacting proteins. Finally, PABPN1 also plays an important role in the transcription of muscle-specific genes, which could explain why other tissues are unaffected [89].

In contrast, HES represents a severe hematologic disorder with sustained overproduction of eosinophils in the bone marrow, eosinophilia, tissue infiltration, and organ damage. In this case, a DNA rearrangement involving a chromosomal deletion of 800 kb and fusion of the hFip1 and PDGFR α genes is the underlying cause of this syndrome [29]. The corresponding chimeric protein, hFip1-PDGFR α , contains the N-terminus of hFip1, an integral subunit of the CPSF complex stimulating 3' end processing [86], and the C-terminal kinase domain of PDGFR α . The expression of the hFip1-PDGFR α fusion protein in hematopoietic cells constitutively activates the PDGFR α kinase and transforms cells. As for OPMD, the resulting phenotype might further be aggravated by an interference with canonical 3' end processing, although this has not been explored in immediate detail. Surprisingly, these examples are currently the only two reported genomic perturbations, which affect processing factors and leading to an overt phenotype. This either reflects a negative selective pressure, an unexpectedly high structural and/or functional flexibility, or the redundancy of some of those factors.

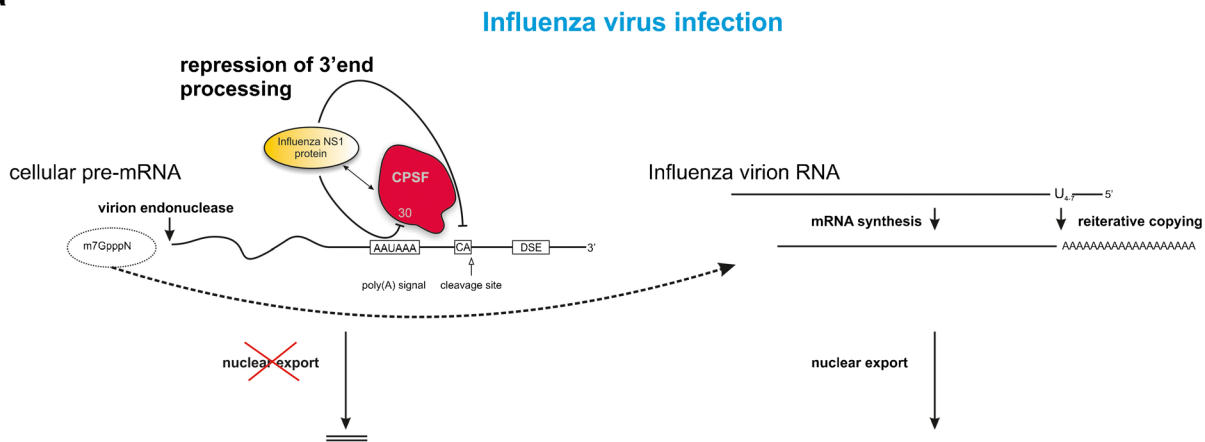
Regulated cleavage and polyadenylation

The complexity of the macromolecular processing machinery in general and different co-factor requirements in particular offer various opportunities for a dynamic regulation of cleavage and polyadenylation ([34] and references therein). Presumably, this plays an important role in the context of physiological adaptation of gene expression but can also result in profound disease phenotypes. In analogy to the regulation of splicing, the decision how efficiently a PAS is used requires RNA sequence elements and protein regulators. Influenza A virus infections provide an interesting example of how the disturbance of protein interactions within the polyadenylation machinery can influence processing: The influenza A NS1 protein is one of the most abundant proteins synthesized in infected cells [97]. It regulates several posttranscriptional processing steps [55, 106] but also interacts with the cellular 30 kDa subunit of CPSF. This sequesters CPSF30 and thereby inhibits 3' end cleavage and polyadenylation of the host pre-mRNAs by preventing the binding of the CPSF complex to the RNA substrate [133] (Fig. 2a). Interestingly, NS1 also targets PABPN1, which inhibits the processive synthesis of long poly(A) tails catalyzed by PAP [25]. As mRNA processing represents a prerequisite for cytoplasmic export [169], the uncleaved host pre-mRNAs are retained in the nucleus, while viral RNAs are still exported. Thus, by interfering with the activity of an essential 3' end processing factor, influenza has devised an efficient way to shut off cellular gene expression and to facilitate viral gene expression. This and other examples [123, 124, 153] illustrate that cleavage and polyadenylation can be regulated by relative simple antagonistic molecular interactions.

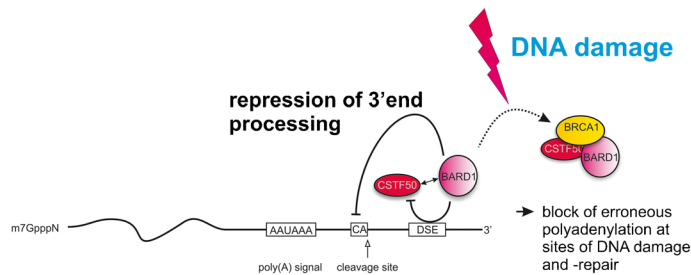
However, an inhibition of 3' end cleavage and polyadenylation can also occur as a result of endogenous cell-intrinsic mechanisms. One prominent example is the BRCA1-associated protein BARD1, which establishes a causal link between mRNA 3' end processing and tumor suppression [90]. BARD1 senses sites of DNA damage and repair and physically interacts with CSTF-50 (Fig. 2b). Challenging cells with DNA-damaging agents transiently inhibit 3' end formation by enhanced formation of CSTF/BARD1/BRCA1 complexes. Furthermore, a tumor-associated germline mutation in BARD1 (Gln564His) decreases its affinity for CSTF-50 and renders the protein inactive in polyadenylation inhibition. The BARD1-mediated inhibition of polyadenylation may thus prevent inappropriate RNA processing during transcription of damaged DNA loci.

Competitive protein interactions modulating the efficiency of cleavage and polyadenylation can be regulated by specific signaling pathways. This has first been demonstrated for the prothrombin (F2) pre-mRNA in which processing relies on a highly conserved upstream sequence element (USE) (Fig. 1a) [31]. In this example, stress conditions that activate p38

a

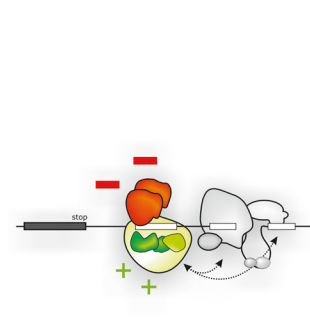


b



c

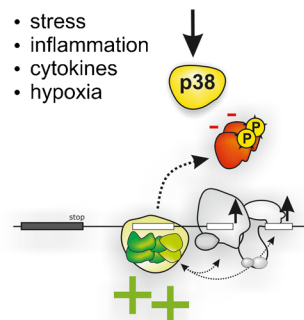
normal state



repressed mRNA 3' end processing

external stimuli

- stress
- inflammation
- cytokines
- hypoxia



enhanced mRNA 3' end processing

d

unstressed



heat shock



MAPK signaling up-modulate components of the 3' end processing apparatus and phosphorylate the RNA-binding proteins FBP2 and FBP3 (Fig. 2c red complex). Normally, these proteins bind to the USE and inhibit 3' end processing. Upon phosphorylation, they dissociate from the USE, making it accessible to proteins that stimulate 3' end processing [36]. These findings have important implications: deregulated F2 expression plays a crucial role in the pathogenesis of thrombophilia but also in other conditions linking p38 MAPK activation with aberrant cellular processes such as tumorigenesis [35]. It is worth

noting that the USE motif constitutes a highly conserved nonameric sequence element, which can be found in many genes including MYC (among other key players with a role in tumorigenesis [36]). Thus, this regulatory principle might account for a plethora of gene functions. From these findings, regulated 3' end processing emerged as a key mechanism of gene regulation with broad biological and medical implications. It also provides a first example how the basal 3' end processing apparatus is “wired” to signaling pathways allowing a dynamic adaptation of the 3' end cleavage and

Fig. 2 Regulated 3' end processing in disease. **a** In influenza A virus-infected cells, the highly abundant NS1 protein interacts with the cellular 30 kDa subunit of CPSF and PABPN1 (not shown) [25]. This prevents binding of the CPSF complex to its RNA substrates and selectively inhibits 3' end processing and nuclear export of host pre-mRNAs. In contrast, the 3' terminal poly(A) sequence on viral mRNAs is produced by the viral transcriptase, which reiteratively copies a stretch of four to seven uridines in the virion RNA templates. In addition, an endonuclease intrinsic to the viral polymerase cleaves cellular capped RNAs to generate capped fragments that serve as primers for the viral mRNA synthesis ("cap-snatching mechanism" [144]). Thus, by interfering with the activity of an essential 3' end processing factor, influenza has devised an efficient way to shut off cellular gene expression and to facilitate viral gene expression [133]. **b** The BRCA1-associated protein BARD1 physically interacts with CSTF-50, thereby repressing the polyadenylation machinery [90]. Both, BARD1 and CSTF-50, also interact with POL2 (not shown), and BARD1 has senses sites of DNA damage and repair. The BARD1-mediated inhibition of polyadenylation thus prevents inappropriate RNA processing during transcription at such compromised sites. Challenging cells with DNA-damaging agents results in a transient inhibition of 3' end formation by enhanced formation of a CSTF/BARD1/BRCA1 complex. A tumor-associated germline mutation in BARD1 decreases its affinity to CSTF-50 and renders the protein inactive in polyadenylation inhibition. These findings link 3' end RNA processing with DNA repair, and loss of wild-type BARD1 could therefore lead to defective control of gene expression as a result of inappropriate polyadenylation. **c** In the human prothrombin (F2) mRNA, the efficiency of 3' end processing is regulated by engagement of a highly conserved USE [36]. After induction of p38 MAPK signaling the USE-RNP architecture changes [31]; inhibitory proteins binding to this element (*red*) are phosphorylated and dissociate from the USE complex while stimulatory 3' end processing components (*green*) are more abundantly loaded onto the USE motif. This together with an induction of the canonical 3' end processing machinery promotes 3' end formation, resulting in higher level of F2 mRNA and protein. This process is believed to play an important role in the deregulation of blood coagulation during septicemia but also in processes such as tumor invasion. **d** The poly(A) polymerase (PAP) that catalyzes the formation of the poly(A) tail can be modified by the poly(ADP-ribose) polymerase 1 (PARP1). This regulates its activity both in vitro and in vivo. During heat shock, PARP1 binds to and modifies PAP leading to inhibition of polyadenylation in a PARP1-dependent manner. The inhibition reflects a reduced RNA binding affinity of PARylated PAP and decreased PAP association with non-heat shock protein-encoding genes [44]. Interestingly, this example also suggests that there must be gene-specific regulatory mechanisms that allow selective gene expression even in conditions, in which PAP as a central enzyme is posttranslationally modified to overcome the initial eliciting event (for further examples see [72])

polyadenylation efficiency. In analogy to other mechanisms (i.e., splicing), this also exemplifies how accessory sequence elements confer specificity to this type of gene regulation.

Finally, another important example establishing the role of posttranslational modifications as a critical element for regulation of 3' end processing is shown for the poly(A) polymerase [44]. This enzyme catalyzing the formation of the poly(A) tail can be posttranslationally modified by the poly(ADP-ribose) polymerase 1 (PARP1) leading to a poly(ADP-ribosylation), which inhibits the PAP activity (Fig. 2d). The physiological importance of this mechanism is shown in the context of heat shock during which PARP1 inhibits

polyadenylation of non-heat shock protein-encoding genes, while polyadenylation of *hsp* transcripts remains unaltered. Thus, a PARP1-mediated modification of PAP has evolved as an effective mechanism for a differential regulation of polyadenylation during thermal stress. Although not fully elucidated, this example also suggests that there must be gene-specific regulatory mechanisms which allow selective gene expression even in conditions, where PAP as a central enzyme is posttranslationally modified [44].

These and other examples illustrate that complex molecular mechanisms have evolved to control and regulate mRNA 3' end processing at (a) defined PAS(s) to eventually execute specific cellular programs. Although not yet explored in further detail, analogous mechanisms might also come into play for the dynamic regulation at alternative ("competing") PASs (next section).

Variations at the transcriptome 3' end—when processing gets alternative

With the emergence of RNA sequencing (RNA-Seq) technologies, it became clear that the transcriptome is enormously diversified at the 3' end [39]. Approximately up to 70 % of the transcriptome is affected by a mechanism widely referred to as "alternative 3' end cleavage and polyadenylation" (APA) [92]. As highlighted above, it regulates numerous genes during the stress response or after T and B cell activation, during differentiation and dedifferentiation, and in various processes linked to tumor progression (detailed below). These findings are in line with earlier observations that alternative PAS selection represents an important and evolutionary conserved regulatory mechanism for spatial (tissue specificity [53, 67, 105, 107]) and temporal control of gene expression (i.e., immunoglobulin class-switch [3, 30, 47, 48, 147, 170, 171]).

The current understanding of how APA is mechanistically controlled is subject of several recent review articles [51, 63, 74, 108, 110, 159, 161, 174]. Although great steps towards a better understanding of APA have been taken, many facets are still enigmatic. Following from above, and possibly despite the fact that APA is widespread, the existence of a unique (and universal?) APA-regulating mechanism is unlikely: In brief, APA can be regulated on the level of mRNA 3' end processing ("direct/true APA") by various *cis*- and *trans*-acting determinants among which (1) the intrinsic strength of sequence elements, (2) the concentration or activity of polyadenylation factors, and/or (3) tissue- or stage-specific regulatory factors are the most important key players [7] (Fig. 3). Yet a significant proportion of APA events is coupled to alternative splicing; PASs located in intronic or alternatively spliced exonic regions can be activated and thus the modulation of splicing can directly influence PAS choice ("splicing coupled APA"; see below). Further, as can be inferred from the coupling of

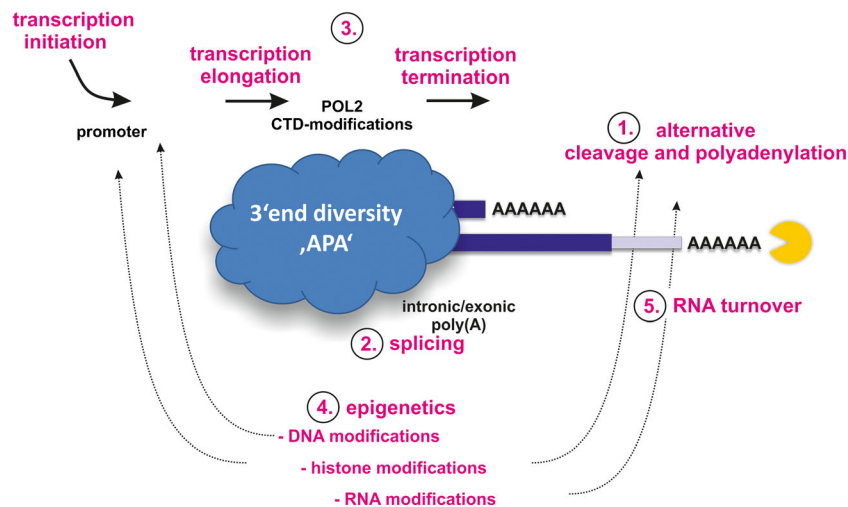


Fig. 3 Mechanisms involved in the regulation of the transcriptome 3' end diversity. Alternative 3' end cleavage and polyadenylation (APA; two transcript isoforms shown) can be regulated (1) on the level of mRNA 3' end processing (“direct/true APA”), through (2) alternative splicing via the in- or exclusion of PASs upon intron retention or exon skipping (“splicing coupled APA”) [183], by (3) transcriptional activities (transcription initiation, elongation, or termination; “kinetic coupling”), or (4) as a result of epigenetic regulation (i.e., through histone or DNA modifications; “epigenetic APA”). Ultimately, changing APA profiles can

also be caused by (5) differential RNA turnover of individual mRNA isoforms, which have been cleaved and polyadenylated at alternative PAS (“faux/indirect APA”). Of note, another interesting mechanistic combination of both splicing regulation and concomitant transcriptional activities is executed via a U1 snRNP-mediated mechanism termed “telescripting” that protects pre-mRNAs from drastic premature termination by cleavage and polyadenylation [11, 83] (APA = alternative cleavage and polyadenylation; CTD = C-terminal domain; POL2 = RNA polymerase II)

cleavage and polyadenylation with transcription (Fig. 1), transcription initiation, the kinetics of transcription itself as well as transcription termination can control PAS choice (“kinetic coupling”; see below). In addition, epigenetic regulation (i.e., through histone or DNA modifications) as well as RNA modifications can determine where and to what extent individual pre-mRNAs are eventually cleaved and polyadenylated (“epigenetic APA”) [57, 87, 102, 167, 182, 190]. Ultimately, it needs to be noted that changing APA phenotypes can also be caused by regulated RNA decay of individual mRNA isoforms, which have been constitutively cleaved and polyadenylated at alternative PAS (“faux/indirect APA”). Thus, dissecting the underlying mechanisms of regulated transcriptome 3' end diversity is complex and has to take into account various variables that importantly can at the same time be cause and consequence of APA (Figs. 3 and 4).

Before moving on to a more universal role of APA in health and disease, we will briefly illustrate a few important physiological aspects, which deserve special attention.

APA in differentiation and development

Recent studies based on high throughput analyses have revealed that APA is highly regulated during development ([8, 54, 76, 79, 152, 158] and references therein). Interestingly there is a correlation between the proliferation status and the global APA patterns (Fig. 4a). Proliferating cells tend to use upstream (“proximal”) PASs and produce mRNAs with

shorter 3'UTRs, while quiescent/differentiated cells favor downstream (“distal”) PASs and produce mRNAs with longer 3'UTRs [152]. Similar observations have been made for mouse development and the differentiation of ECSs into neurons and other functions [16, 78, 158]. In contrast, during somatic reprogramming [79] or tumorigenesis, proximal PASs are favored leading to shorter 3'UTRs [121]. In some of those cases, mRNAs with shorter 3'UTRs tend to be more stable [71, 162] or globally elevated [40] eventually leading to higher protein output [121, 152]. Breaking it down to individual transcripts, the consequences of APA however can be complex: APA transcript isoforms of the same gene can encode different proteins and/or change the 3'UTR properties, leading to the inclusion or exclusion of mRNA stabilizing or destabilizing elements, miRNA target sites, or result in different translation efficiencies or subcellular localization (Fig. 4a, detailed below). Thus, although APA is widespread in processes such as differentiation, dedifferentiation, and development, global equations of how the overall trend of APA directionality affects the fate of the respective transcript isoforms are difficult. This is supported by studies, which failed to detect a straightforward correlation between APA and mRNA stability or protein output [50, 58, 64, 130, 152, 166]. Accordingly, deciphering the downstream functional consequences of APA in the context of cellular programs and to understand whether APA is the cause or consequence of complex biological programs represents a major challenge. Possibly APA coordinately regulates post-transcriptional regulons (“RNA operons”) driving specific cellular programs.

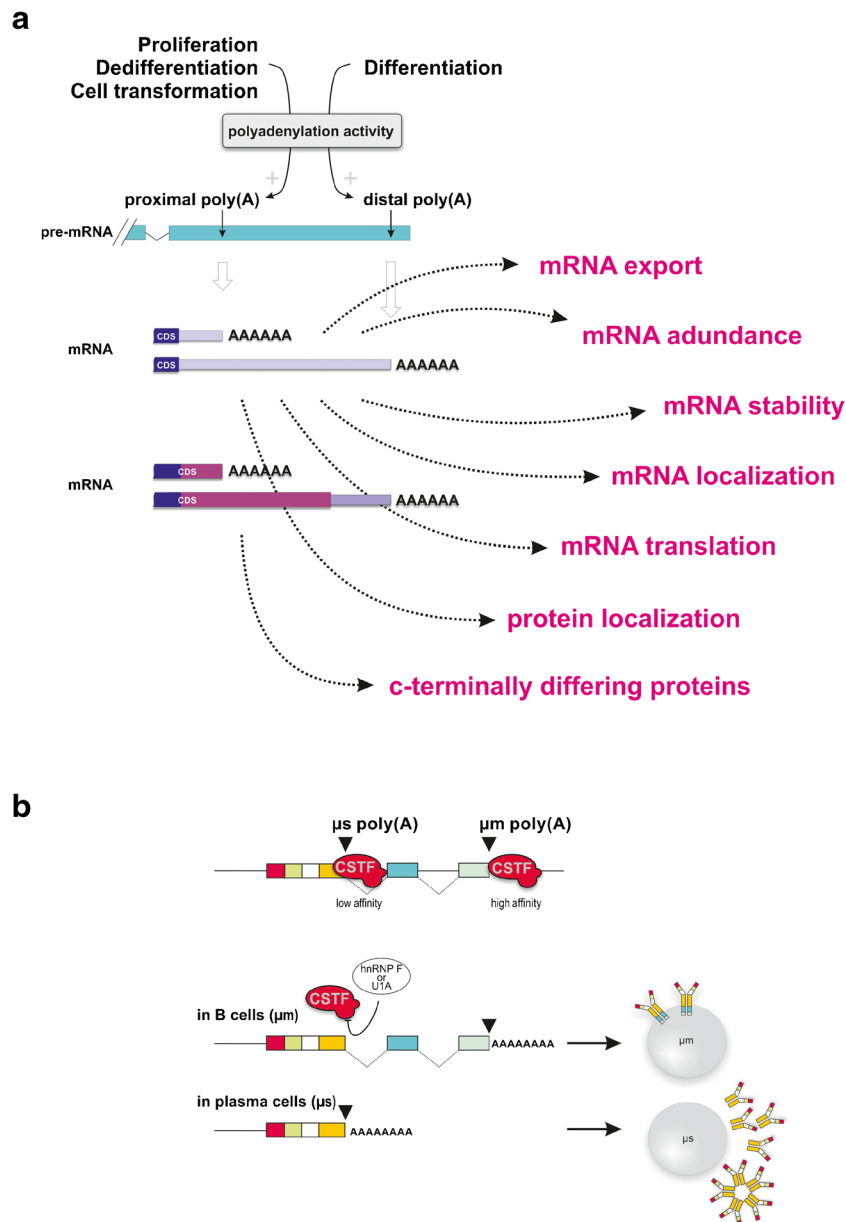


Fig. 4 Alternative 3' end processing modulates the temporal and spatial diversity of gene expression. **a** About half of the human pre-mRNAs contain (multiple) alternative poly(A) signals (PAS), which are mostly located as tandem arrays within the 3'-UTRs [173, 194]. A smaller set of pre-mRNAs bears alternative PAS within intronic or exonic regions. In both cases, endogenous and exogenous factors can modulate pre-mRNA PAS selection by interfering with constitutive and/or auxiliary 3' end processing factors/subunits. This results in various polyadenylated mRNA isoforms that either code for identical (tandem terminal PASs) or C-terminally modified (internal PASs) proteins. Furthermore, alternatively 3' end processed mRNAs can display different 3'UTR properties. This can affect various aspects of mRNA and/or protein fates (export, abundance, stability, localization, transport, and translation), i.e., via the in- or exclusion of regulatory elements such as microRNA target sites or binding sites for RNA binding proteins with important roles in Mendelian disease [22]. Very commonly, cellular states

associated with enhanced proliferation, dedifferentiation, or cell transformation tend to produce shorter mRNA variants (cleavage and polyadenylation occurring at the proximal PAS), while differentiation shifts cleavage and polyadenylation towards production of longer mRNA variants (processed at the distal PAS). **b** During B cell differentiation, alternative PAS selection effects a switch of the IgM heavy chain expression from a membrane bound form (μm) to the secreted form (μs). In this example, CSTF-64 binding to the PAS of the RNA giving rise to secreted IgM (μs) is favored either by high CSTF-64 concentrations [171] or under conditions of low hnRNP F and/or low U1A concentrations [139, 177] in plasma cells (*lower lane*). In contrast, the high affinity site of the membrane bound form (μm) is used in B cells (*upper lane*) where the CSTF-64 concentration is low or when high concentrations of U1A and/or hnRNP F inhibit CSTF-64 binding to the secretory μs -specific PAS (*boxes* indicate exons, simplified representation)

For example, hFip1 (an integral part of the CPSF complex) has recently been shown to control embryonic stem cell-specific

APA profiles to ensure the optimal expression of a specific set of genes, including critical self-renewal factors, in the cell fate

specification [94]. Yet, we have just begun to decipher the resulting consequences, and further studies are needed to expand our understanding of the resulting consequences. In contrast, the impact of APA on the expression and function of individual genes is far better understood.

Role of APA for individual genes

The historically eldest and perhaps the most thoroughly studied example illustrating the importance of APA and also shedding light on the underlying regulatory principles is the regulation of IgM heavy chain expression during B cell differentiation (Fig. 4b; [3, 47, 147]). In this mRNA, alternative PAS selection is regulated by a modulated recruitment of the CSTF 64-kDa subunit to one of two competing PASs. Upon B cell activation, this switches the IgM heavy chain expression from a membrane bound form (μm) to the secreted form (μs) by activation of an alternative upstream μs -specific PAS in plasma cells [171]. Several mechanisms are possibly contributing including additional modulators (U1A) tightly controlling and ensuring appropriate PAS selection [138, 139]. Although the underlying regulation is further complicated by being coupled to alternative splicing (“splicing coupled APA”), it establishes an important direct functional link between dynamically regulated APA and a physiological process.

A similar mechanism underlies the regulated expression of the transcription factor NF-ATc during T cell differentiation [27]. Two longer isoforms of NF-ATc mRNA are synthesized in naïve T cells, whereas a shorter isoform is expressed in effector cells. The switch is mediated by activation of a proximal PAS by up-regulation of CSTF 64 kDa subunit, which occurs upon T cell stimulation (“direct APA”). In this context, it is interesting to note that LPS stimulation increases CSTF 64 expression in macrophages, which in turn regulates APA of several mRNAs [157]. These examples illustrate that the modulated expression of individual 3' end processing factors can directly—or in concert with auxiliary (inhibitory or stimulatory) factors—change the production of APA isoforms and thereby drive important cellular functions.

Inferring from the tight coupling of transcription with processing [115, 131], the kinetics of POL2 play an important role in PAS choice (“kinetic coupling”). In *Drosophila*, *polo* is a cell cycle gene, which uses two PAS in the 3' UTR to produce alternative messenger RNAs that differ in their 3' UTR length. By using a mutant *Drosophila* strain with a lower transcriptional elongation rate, it was shown that transcription kinetics can determine alternative PAS selection. Although only one gene is affected, the physiological consequences of incorrect *polo* PAS choice are detrimental; transgenic flies lacking the distal poly(A) signal cannot produce the longer transcript and die at the pupa stage due to a failure in the proliferation of the precursor cells of the abdomen [140].

Along these lines also, transcription elongation factors can direct alternative RNA processing and thereby control important cellular functions such as the immunoglobulin secretion in plasma cells [117].

Another interesting example is the brain-derived neurotrophic factor (BDNF), which is encoded by two transcripts with either short or long 3' UTRs. The physiological significance of the two mRNA isoforms encoding the same protein has been unknown until it could be demonstrated that the short and long 3' UTR BDNF mRNAs are involved in different cellular functions. The short 3' UTR mRNAs are restricted to somata, whereas the long 3' UTR mRNAs are also localized in dendrites. In a mouse mutant where the long 3' UTR is truncated, dendritic targeting of BDNF mRNAs is impaired, resulting in low level BDNF in hippocampal dendrites, a selective impairment in long-term potentiation in dendrites, while somata of hippocampal neurons remained normal. These results provide insights into local and dendritic actions of BDNF and reveal APA for a differential regulation of sub-cellular functions of proteins [4] with important medical implications [103].

Further examples documenting the biological consequences of APA are represented by the regulated expression of a truncated form of glutamyl-prolyl tRNA synthetase (EPRS), which as a “gamma-interferon-activated inhibitor of translation” (GAIT) constituent controls the translation of GAIT target transcripts such as the VEGF-A [195]. Furthermore, recent studies demonstrated APA's potential to differentially regulate the localization of membrane proteins by a trafficking mechanism involving the CD47 3'UTR as a scaffold [12]. Altogether, these and other examples illustrate the physiological importance of regulated mRNA 3' end processing as a mechanism controlling a wide spectrum of cellular functions.

APA in human disease

Aberrant APA profiles are associated with a variety of human disorders ([8, 40, 119, 121, 130] and references therein). Most importantly, the strong prevalence of APA regulation in physiological processes such as differentiation and development (see above) is also reflected in situations where these processes are typically dysregulated. The most prototypical example for this is uncontrolled cellular proliferation in the course of cancer development. Accordingly, a widespread increase in the use of proximal PASs has been observed in various cancer cells [121]. In this context, the shorter mRNA isoforms showed an increased stability and typically produced more protein, in part through the loss of microRNA-mediated repression. Interestingly, switching to shorter 3'UTRs also allowed proto-oncogenes to escape from inhibition by miRNAs, thereby resulting in oncogene activation in the

absence of genetic alterations. Global induction of proximal PAS usage has consistently been observed in several studies ever since ([2, 58, 104, 119, 130, 162, 191] and references therein). This phenotype however does not apply to all tumor types [58], and occasionally the correlation between cancer progression and 3'UTR shortening appears to be more complex [49, 58, 130]. Interestingly, the determination of (selected) APA profiles has recently proven to be of prognostic significance [99, 119, 186, 191]. Yet, also other disorders such as endocrine [146] or cardiovascular disease [164] are associated with a widespread regulation of APA.

Previously, PABPN1 has been identified to represent a potent modulator of APA by inhibiting processing at respective PASs [77]. In the context of OPMD (see above), these data may also imply that OPMD is associated with misregulated APA, which results in unbalanced formation of alternative mRNA 3' ends. They also predict that OPMD symptoms in humans may be to a certain degree a result of aberrant gene expression due to a change in 3' end formation. Yet, so far, this has not been tested directly.

Interestingly, also pathological changes occurring in the context of myotonic dystrophy may be attributable to specific alterations in 3'UTR structures and subsequent changes in RNA localization and/or protein isoforms and levels [8]. Although primarily demonstrated in mice, complementary analysis of samples of human origin strongly suggests APA to represent a (additional) molecular mechanism underlying muscular dystrophy.

Given the high prevalence of APA as a pervasive gene regulatory mechanism in various physiologically relevant processes, it is likely that its misregulation will be discovered in the context of various other pathological conditions [172]. Yet the extent to which the vast majority of all reported global alterations at the mRNA 3' end represent driver or passengers of human disorders remains an open question. Furthermore, it is, for example, conceivable that the observed APA pattern changes could equally reflect compensatory activities reestablishing the cellular homeostasis in response to disease-triggering events.

Potentially interesting evidence for a causal relationship between APA and a resulting disorder is represented by a polymorphism in the interferon regulatory factor (IRF) 5 gene predisposing to systemic lupus erythematosus (SLE) [61, 62]. This newly identified polymorphism creates a functional polyadenylation site resulting in an increased expression of a transcript variant containing a shorter 3'UTR. Interestingly, the expression levels of transcript variants with the shorter or longer 3'UTRs appeared to be inversely correlated. Thereby, it contributes to a misregulation of interferon signaling, a critical constituent in the pathogenesis or progression of SLE. Yet, this is obviously one of the “simpler” examples in which APA (of one gene) is altered in *cis*. Accordingly, the nature and functional consequences of global APA regulation in the

context of human pathologies remain subject to future investigations.

APA in molecular diagnostics

The advent of high-throughput sequencing technologies has significantly promoted the elucidation of disease mechanisms and equipped us with novel diagnostic opportunities. As for alternative splicing, identifying APA patterns can likely have wide diagnostic implications [38, 84, 181]. With the evolution of numerous protocols that rely on 3' end sequencing technologies [42, 56, 58, 73, 76, 77, 104, 113, 116, 137, 158, 183, 188, 196], the determination of APA isoforms can meanwhile be carried out in a reliable and straightforward manner (Fig. 5). Unlike gene expression profiling based on arrays or full RNA-Seq, which can be dramatically confounded by the way of normalization, the relative proportion of APA isoforms is normally internally controlled thus resulting in relatively robust results. While contamination of the analyzed specimen by other cell populations is another inevitable and very common confounder in gene expression analysis (i.e., when revealing the signature of a tumor which is infiltrated by immune cells), APA patterns are likely to be more tissue specific [42] and were found to differ according to tissue type, developmental stage, genotype, or cancer subtype [78, 79, 92, 111, 152, 162, 183, 197]. Cataloging these (tissue) specific patterns might therefore allow us to subtract “contaminating” APA signatures from APA signatures of specific disordered tissues of interest. Thus, the determination of APA patterns may open up novel diagnostic avenues which up to this point have turned out to represent challenging aspects of “conventional” gene expression profiling. Finally, characterizing tissue-specific APA signatures per se may be of immediate diagnostic value, i.e., for tracing back and identifying the origin of a given disease lesion (for instance in cases of *cancer of unknown primary*, CUPs).

With the advent of high-throughput analyses, the bioinformatical workload has increased dramatically. In contrast to total RNA-Seq, the sequencing restricted to the transcriptome 3' end directly uncovers the variability and perturbation occurring at the mRNA 3' end (Fig. 5). This has several advantages. Firstly, it drastically reduces the bioinformatical workload. Furthermore, these data are typically not confounded by other variables that complicate the bioinformatical processing of the data (such as alternative splicing). Finally, restricting the sequencing to the last (approximately) 30 nucleotides of the transcriptome opens up interesting (and first and foremost cost-effective) opportunities for multiplexing—while still keeping a high coverage for a reliable analysis. In depth APA profile studies have recently revealed “aberrant” APA signatures to be associated with more aggressive tumor phenotypes in cancer patients and thereby provided the proof-

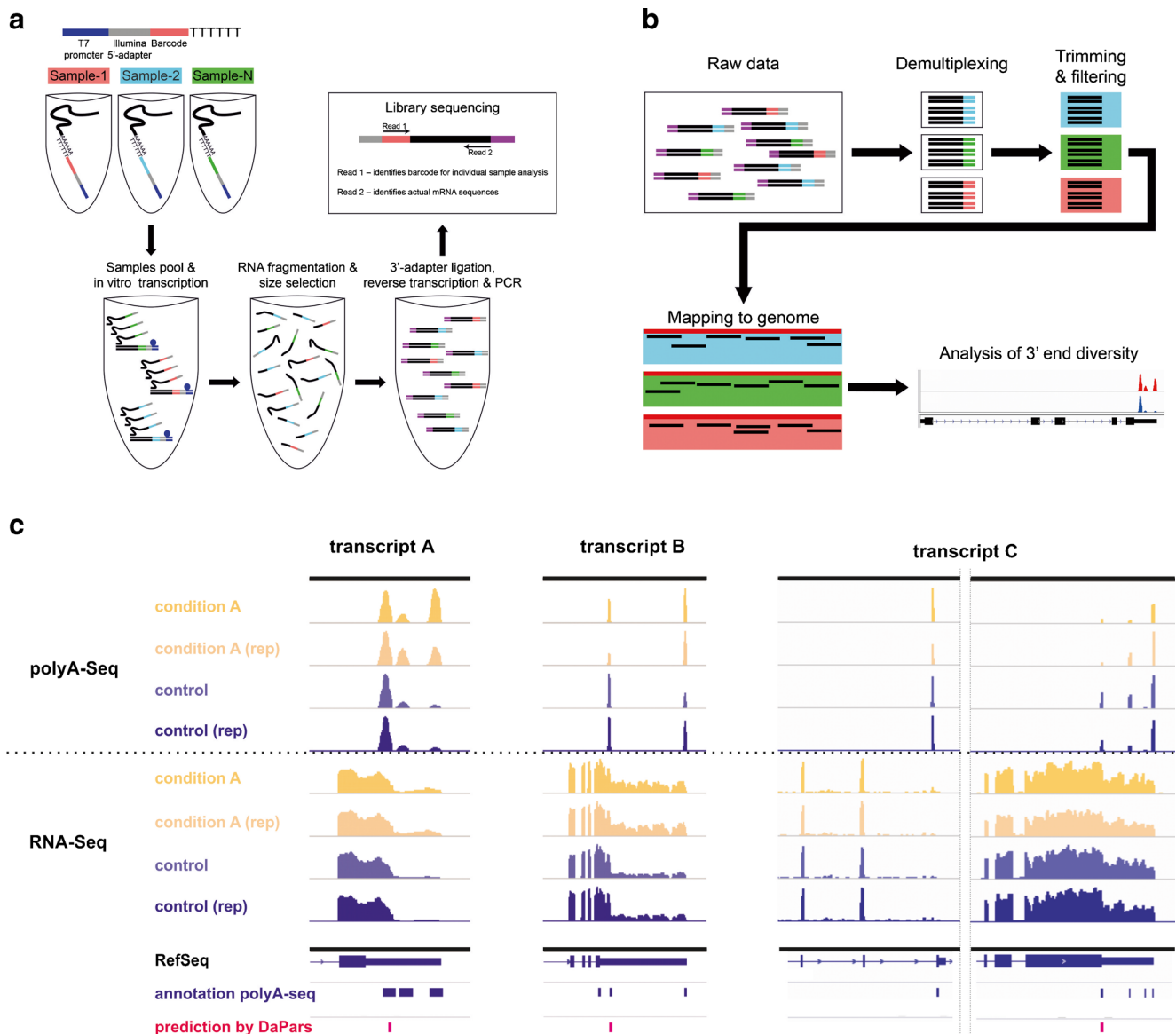


Fig. 5 High throughput sequencing for the (multiplexed) definition of mRNA 3' end diversity of gene expression. **a** Principle for the definition of the transcriptome 3' end ("polyA-Seq") via (multiplexed) high-throughput sequencing based on oligo(dT) primed, barcoded cDNA library generation. **b** Workflow showing the bioinformatical processing of the sequencing reads involving demultiplexing, filtering, and mapping for the visualization of 3' end seq reads. **c** Visualization of the

transcriptome 3' end (three selected examples) by polyA-Seq compared to RNA-Seq and definition of the mRNA 3' end by DaPars on the basis of RNA-Seq data [191]. Shown are differential changes of the mRNA 3' end signature comparing two conditions ("condition A" and "control," each in biological replicates (Danckwardt lab, unpublished)). The lower panel illustrates the respective RefSeq 3'UTR and the annotation of the polyA-Seq data in comparison with a differential use of PAS in DaPars

of-concept that such a determination can reveal prognostic signatures [119, 191]. Yet, applying novel bioinformatical analysis (DaPars), APA patterns can also be extracted from preexisting transcriptome wide sequencing data [191]. Although this takes advantage of the fact that RNA-Seq data is already available for numerous tissue specimens, this technique has the limitation that it is primarily suited to detect alternative 3'UTR events, while APA events, which are located within the coding region, or alternatively spliced introns (internal APA) rather remain obscure. Compared to 3' end sequencing technologies, this algorithm requires complex

bioinformatical calculations and typically allows a less "intuitive" identification of the mRNA 3' end (Fig. 5c, compare polyA-Seq with DaPars).

It remains to be observed in which disease conditions and to what extent the analysis of APA signatures could further improve diagnostic strategies and possibly allow detecting biological aberrations with higher sensitivity and specificity. Interestingly, selected APA events can confer strong prognostic power beyond common clinical and molecular variables, suggesting their potential as novel prognostic biomarkers [191]. Thus, it will be interesting to see how the determination

of APA patterns may evolve as a potentially new biomarker in the future. This could advance diagnostic strategies for a more thorough understanding of underlying disease mechanisms as well as for a reliable prognostic and possibly therapeutic stratification.

Ultimately, ongoing genome sequencing activities will most likely grant us further insights into genomic variations resulting in gene-specific perturbation of APA isoforms with possible detrimental functional consequences. Unlike global aberration in *trans* (i.e., as a result of a change of the abundance of one processing factor or regulatory protein), the cause-consequence relationship in this kind of setting is substantively clearer. Further such changes may be directly accessible for specific, targeted therapeutic approaches.

Targeting mRNA 3' end formation as a novel therapy

We have seen that the determination of global APA patterns can have important diagnostic and even prognostic implications. The therapeutic significance of APA will depend on it being cause, consequence, or simply a coincidental epiphenomenon of the underlying disease.

However, even the latter two conditions do not necessarily preclude the possibility of regulating APA as a therapeutically meaningful approach. Various disorders are associated with drastic APA changes (see above), and it is difficult to imagine that all observed APA patterns are biologically “silent” and consequently do not affect a potential phenotype.

Although still on an experimental level, in principle, strategies to interfere with 3' end processing are available. This encompasses both unspecific as well as target-specific strategies. For example, as shown for the regulation of splicing, antisense oligonucleotides (ASO) inhibiting U1 binding can be used to specifically promote intronic alternative polyadenylation [179]. Although the experience concerning therapeutic targeting of splicing is far more advanced, the general proof-of-concept of targeting specific polyadenylation sites for redirection of processing based on analogous approaches has been made. This includes the use of ASOs [179] and siRNAs [178] as well as modified U1 snRNP, which interacts with a target gene upstream of its PAS to regulate gene expression [14].

Yet, also other strategies might come into play as well. We have seen that APA is influenced by various other cellular processes controlling gene expression (Figs. 1 and 3) including the velocity/kinetics of POL2 [140]. Although presumably unspecific, the interference with POL2 processivity at various stages of transcription [81] may potentially regulate APA [46]. In fact, numerous anticancer drugs regulate in the one or the other way the processivity of POL2 (such as doxorubicin or camptothecin). Apart from this, the C-terminal domain of POL2 is subject to extensive posttranslational phosphorylation,

which influences co-transcriptional events including splicing, transcription termination, and 3' end processing [68]. Interestingly, although Ser 2 and Ser 5 phosphorylations of the CTD are by far the most studied posttranslational modifications, the way in which the phosphorylation pattern itself or potentially even other posttranslational modifications might affect the loading and delivery of processing factors to their ultimate destination is yet to be elucidated.

Furthermore, 3' end processing is tightly bound to splicing, and a significant proportion of APA events per se occur concurrently with alternative splicing. Therefore, virtually all therapeutical approaches currently tested for manipulating splicing [163] may—in the one or the other instance—help reverting “disordered” APA phenotypes as well. Ultimately, we have obtained first evidence how extracellular signals influence the basal 3' end processing machinery [31]. Generally, this and other examples connecting posttranslational modifications with the regulation of 3' end processing as shown for PAP [44] or the modifications of POL2 CTD (see above) may lead new ways towards targeting signaling components for regulation of the transcriptome 3' end diversity. Finally, by means of the ongoing research elucidating how epigenetic modifications can control APA switches ([102, 182, 190] and references therein), it is tempting to speculate that the manipulation of these pathways may eventually be translated into the clinical context.

Conclusions and perspectives

The 3'UTR has emerged as a hotspot for posttranscriptional gene regulation, controlling important cellular functions such as morphogenesis, cell differentiation, metabolism, cell proliferation, and other processes by controlling mRNA translation, stability, localization, as well as 3' end processing. More recently, APA has been found to represent an important layer of posttranscriptional gene regulation, which in turn can influence the RNA fate and/or regulate the protein output quantitatively or qualitatively, and thereby steer important cellular programs. Interestingly, we have seen distorted APA signatures being associated with a variety of disorders. Presumably, more of these patterns will be discovered in the context of various other pathological conditions in the near future.

From a medical perspective, the most intriguing question relates to the extent to which APA represents a driver or passenger of human disorders. For individual genes harboring mutations that perturb alternative cleavage and polyadenylation of their own transcripts, the cause-consequence relationship is relatively simple. However, the causal contribution of a widespread APA (de)regulation in the context of human pathologies is still unclear. The broad biological importance of APA renders the possibility of global changes of the transcriptome 3' end

(associated with human pathologies) to have any phenotypic consequences unlikely. Even without being disease eliciting directly, these changes may aggravate underlying pathologies. Yet, they could equally represent compensatory activities for reestablishing the cellular homeostasis in response to disease-triggering events.

Thus, further studies are required to decipher the functional contribution of regulated APA in the context of human pathologies in order to determine whether APA can serve as a meaningful therapeutic target. Defining key components directing APA and dissecting their functional hierarchy thus represent critical aspects that influence the conceptual tractability—apart from all practical possibilities/opportunities. Uncoupled from these challenging aspects, exiting first steps towards APA serving as a potentially novel biomarker are taken. It will be interesting to see how these encouraging findings will further translate into a clinical setting and whether they will become part of “routine” molecular diagnostics allowing prognostic and/or therapeutic stratifications in the (near) future.

Acknowledgments We thank members of the Danckwardt lab for helpful discussions. We apologize to all colleagues whose work could not be discussed or cited here because of space constraints.

Work in the laboratory of SD is supported by the DFG (DA 1189/2-1), by the GRK 1591, by the Federal Ministry of Education and Research (BMBF 01EO1003), by the Hella Bühler Prize for Cancer Research, and by the DGKL and the Institute of Clinical Chemistry, University Medical Center Mainz. We thank Vanessa Rau for helpful input and critical comments on the manuscript.

Author contributions AO, YK, and SD selected and analyzed the literature and drafted the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Ahn SH, Kim M, Buratowski S (2004) Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Mol Cell* 13:67–76
- Akman BH, Can T, Erson-Bensan AE (2012) Estrogen-induced upregulation and 3'-UTR shortening of CDC6. *Nucleic Acids Res* 40:10679–10688. doi:10.1093/nar/gks855
- Alt FW, Bothwell ALM, Knapp M, Siden E, Mather E, Koshland M, Baltimore D (1980) Synthesis of secreted and membrane-bound immunoglobulin mu heavy chains is directed by mRNAs that differ at their 3-prime ends. *Cell* 20:293–301
- An JJ, Gharami K, Liao G-Y, Woo NH, Lau AG, Vanevski F, Torre ER, Jones KR, Feng Y, Lu B, Xu B (2008) Distinct role of long 3'UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell* 134:175–187. doi:10.1016/j.cell.2008.05.045
- Awasthi S, Alwine JC (2003) Association of polyadenylation cleavage factor I with U1 snRNP. *RNA* 9:1400–1409
- Balim Z, Kosova B, Falzon K, Bezzina Wettinger S, Colak Y (2003) Budd-Chiari syndrome in a patient heterozygous for the point mutation C20221T of the prothrombin gene. *J Thromb Haemost* 1:852–853
- Barabino SM, Keller W (1999) Last but not least: regulated poly(A) tail formation. *Cell* 99:9–11
- Batra R, Charizanis K, Manchanda M, Mohan A, Li M, Finn DJ, Goodwin M, Zhang C, Sobczak K, Thornton CA, Swanson MS (2014) Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol Cell* 56:311–322. doi:10.1016/j.molcel.2014.08.027
- Beaudoing E, Freier S, Wyatt JR, Claverie JM, Gautheret D (2000) Patterns of variant polyadenylation signal usage in human genes. *Genome Res* 10:1001–1010
- Bennett CL, Brunkow ME, Ramsdell F, O'Brian KC, Zhu Q, Fuleihian RL, Shigeoka AO, Ochs HD, Chance PF (2001) A rare polyadenylation signal mutation of the FOXP3 gene (AAUAAA->AAUGAA) leads to the IPEX syndrome. *Immunogenetics* 53:435–439
- Berg MG, Singh LN, Younis I, Liu Q, Pinto AM, Kaida D, Zhang Z, Cho S, Sherrill-Mix S, Wan L, Dreyfuss G (2012) U1 snRNP determines mRNA length and regulates isoform expression. *Cell* 150:53–64. doi:10.1016/j.cell.2012.05.029
- Berkovits BD, Mayr C (2015) Alternative 3[prime] UTRs act as scaffolds to regulate membrane protein localization. *Nature* 522:363–367. doi:10.1038/nature14321
- Black DL (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* 72:291–336
- Blázquez L, Fortes P (2015) U1 interference (U1i) for antiviral approaches. In: Berkhout B, Ertl CJH, Weinberg SM (eds) Gene therapy for HIV and chronic infections. Springer New York, New York, pp 51–69. doi:10.1007/978-1-4939-2432-5_3
- Blencowe BJ (2006) Alternative splicing: new insights from global analyses. *Cell* 126:37–47. doi:10.1016/j.cell.2006.06.023
- Boutet SC, Cheung TH, Quach NL, Liu L, Prescott SL, Edalati A, Iori K, Rando TA (2012) Alternative polyadenylation mediates microRNA regulation of muscle stem cell function. *Cell Stem Cell* 10:327–336. doi:10.1016/j.stem.2012.01.017
- Brackenridge S, Proudfoot NJ (2000) Recruitment of a basal polyadenylation factor by the upstream sequence element of the human lamin B2 polyadenylation signal. *Mol Cell Biol* 20:2660–2669
- Brais B, Bouchard JP, Xie YG, Rochefort DL, Chretien N, Tome FM, Lafreniere RG, Rommens JM, Uyama E, Nohira O, Blumen S, Korczyn AD, Heutink P, Mathieu J, Duranceau A, Codere F, Fardeau M, Rouleau GA, Korczyn AD (1998) Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat Genet* 18:164–167
- Brown KM, Gilmartin GM (2003) A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. *Mol Cell* 12:1467–1476
- Buratowski S (2005) Connections between mRNA 3' end processing and transcription termination. *Curr Opin Cell Biol* 17:257–261
- Calado A, Tome FM, Brais B, Rouleau GA, Kuhn U, Wahle E, Carmo-Fonseca M (2000) Nuclear inclusions in oculopharyngeal muscular dystrophy consist of poly(A) binding protein 2 aggregates which sequester poly(A) RNA. *Hum Mol Genet* 9:2321–2328

22. Castello A, Fischer B, Hentze MW, Preiss T (2013) RNA-binding proteins in Mendelian disease. *Trends Genet* 29:318–327. doi:10.1016/j.tig.2013.01.004
23. Chan SL, Huppertz I, Yao C, Weng L, Moresco JJ, Yates JR 3rd, Ule J, Manley JL, Shi Y (2014) CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3' processing. *Genes Dev* 28:2370–2380. doi:10.1101/gad.250993.114
24. Chang H, Lim J, Ha M, Kim VN (2014) TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. *Mol Cell* 53:1044–1052. doi:10.1016/j.molcel.2014.02.007
25. Chen Z, Li Y, Krug RM (1999) Influenza A virus NS1 protein targets poly(A)-binding protein II of the cellular 3'-end processing machinery. *EMBO J* 18:2273–2283
26. Chen F, MacDonald CC, Wilusz J (1995) Cleavage site determinants in the mammalian polyadenylation signal. *Nucleic Acids Res* 23:2614–2620
27. Chuvpilo S, Zimmer M, Kerstan A, Glockner J, Avots A, Escher C, Fischer C, Inashkina I, Jankevics E, Berberich-Siebelt F, Schmitt E, Serfling E (1999) Alternative polyadenylation events contribute to the induction of NF-ATc in effector T cells. *Immunity* 10:261–269
28. Colgan DF, Manley JL (1997) Mechanism and regulation of mRNA polyadenylation. *Genes Dev* 11:2755–2766
29. Cools J, DeAngelo DJ, Gotlib J, Stover EH, Legare RD, Cortes J, Kutok J, Clark J, Galinsky I, Griffin JD, Cross NCP, Tefferi A, Malone J, Alam R, Schrier SL, Schmid J, Rose M, Vandenberghe P, Verhoef G, Boogaerts M, Wlodarska I, Kantarjian H, Marynen P, Coutre SE, Stone R, Gilliland DG (2003) A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. *N Engl J Med* 348:1201–1214
30. Cushley W, Coupar BE, Mickelson CA, Williamson AR (1982) A common mechanism for the synthesis of membrane and secreted immunoglobulin alpha, gamma and mu chains. *Nature* 298:77–79
31. Danckwardt S, Gantzer AS, Macher-Goeppinger S, Probst HC, Gentzel M, Wilm M, Grone HJ, Schirmacher P, Hentze MW, Kulozik AE (2011) p38 MAPK controls prothrombin expression by regulated RNA 3' end processing. *Mol Cell* 41:298–310
32. Danckwardt S, Gehring NH, Neu-Yilik G, Hundsdorfer P, Pforsich M, Frede U, Hentze MW, Kulozik AE (2004) The prothrombin 3' end formation signal reveals a unique architecture that is sensitive to thrombophilic gain-of-function mutations. *Blood* 104:428–435
33. Danckwardt S, Hartmann K, Katz B, Hentze M, Levy Y, Eichele R, Deutsch V, Kulozik A, Ben-Tal O (2006) The prothrombin 20209 C>T mutation in Jewish-Moroccan Caucasians: molecular analysis of gain-of-function of 3' end processing. *J Thromb Haemost* 4:1078–1085
34. Danckwardt S, Hentze MW, Kulozik AE (2008) 3' end mRNA processing: molecular mechanisms and implications for health and disease. *EMBO J* 27:482–498
35. Danckwardt S, Hentze MW, Kulozik AE (2013) Pathologies at the nexus of blood coagulation and inflammation: thrombin in hemostasis, cancer and beyond. *J Mol Med* 91:1257–1271
36. Danckwardt S, Kaufmann I, Gentzel M, Foerstner KU, Gantzer AS, Gehring NH, Neu-Yilik G, Bork P, Keller W, Wilm M, Hentze MW, Kulozik AE (2007) Splicing factors stimulate polyadenylation via USEs at non-canonical 3' end formation signals. *EMBO J* 26:2658–2669
37. Darmon SK, Lutz CS (2012) Novel upstream and downstream sequence elements contribute to polyadenylation efficiency. *RNA Biol* 9:1255–1265. doi:10.4161/ma.21957
38. David CJ, Manley JL (2010) Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes Dev* 24:2343
39. de Klerk E, 't Hoen PAC (2015) Alternative mRNA transcription, processing, and translation: insights from RNA sequencing. *Trends Genet* 31:128–139. doi:10.1016/j.tig.2015.01.001
40. de Klerk E, Venema A, Anvar SY, Goeman JJ, Hu O, Trollet C, Dickson G, den Dunnen JT, van der Maarel SM, Raz V, 't Hoen PAC (2012) Poly(A) binding protein nuclear 1 levels affect alternative polyadenylation. *Nucleic Acids Res* 40:9089–9101. doi:10.1093/nar/gks655
41. Decorsière A, Toulas C, Fouque F, Tilkin-Mariamé A-F, Selves J, Guimbaud R, Chipoulet E, Delmas C, Rey J-M, Pujol P, Favre G, Millevoi S, Vagner S (2012) Decreased efficiency of MSH6 mRNA polyadenylation linked to a 20-base-pair duplication in Lynch syndrome families. *Cell Cycle* 11:2578–2580
42. Derti A, Garrett-Engel P, Macisaac KD, Stevens RC, Sriram S, Chen R, Rohl CA, Johnson JM, Babak T (2012) A quantitative atlas of polyadenylation in five mammals. *Genome Res* 22:1173–1183. doi:10.1101/gr.132563.111
43. Di Giammartino DC, Manley JL (2014) New links between mRNA polyadenylation and diverse nuclear pathways. *Mol Cell* 37:644–649. doi:10.14348/molcells.2014.0177
44. Di Giammartino DC, Shi Y, Manley JL (2013) PARP1 represses PAP and inhibits polyadenylation during heat shock. *Mol Cell* 49:7–17. doi:10.1016/j.molcel.2012.11.005
45. Dominski Z, Yang XC, Marzluff WF (2005) The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. *Cell* 123:37–48
46. Dutertre M, Chakrama FZ, Combe E, Desmet FO, Mortada H, Polay Espinoza M, Gratadou L, Auboeuf D (2014) A recently evolved class of alternative 3'-terminal exons involved in cell cycle regulation by topoisomerase inhibitors. *Nat Commun* 5:3395. doi:10.1038/ncomms4395
47. Early P, Rogers J, Davis M, Calame K, Bond M, Wall R, Hood L (1980) Two mRNAs can be produced from a single immunoglobulin mu gene by alternative RNA processing pathways. *Cell* 20:313–319
48. Edwalds-Gilbert G, Veraldi KL, Milcarek C (1997) Alternative poly(A) site selection in complex transcription units: means to an end? *Nucleic Acids Res* 25:2547–2561
49. Elkon R, Drost J, van Haaften G, Jenal M, Schrier M, Oude Vrielink JA, Agami R (2012) E2F mediates enhanced alternative polyadenylation in proliferation. *Genome Biol* 13:R59. doi:10.1186/gb-2012-13-7-r59
50. Elkon R, Drost J, van Haaften G, Jenal M, Schrier M, Vrielink JAO, Agami R (2012) E2F mediates enhanced alternative polyadenylation in proliferation. *Genome Biol* 13:1–15. doi:10.1186/gb-2012-13-7-r59
51. Elkon R, Ugalde AP, Agami R (2013) Alternative cleavage and polyadenylation: extent, regulation and function. *Nat Rev Genet* 14:496–506. doi:10.1038/nrg3482
52. Faustino NA, Cooper TA (2003) Pre-mRNA splicing and human disease. *Genes Dev* 17:419–437
53. Feng L, Sun W, Xia Y, Tang WW, Chanmugam P, Soyoola E, Wilson CB, Hwang D (1993) Cloning two isoforms of rat cyclooxygenase: differential regulation of their expression. *Arch Biochem Biophys* 307:361–368
54. Flavell SW, Kim TK, Gray JM, Harmin DA, Hemberg M, Hong EJ, Markenscoff-Papadimitriou E, Bear DM, Greenberg ME (2008) Genome-wide analysis of MEF2 transcriptional program reveals synaptic target genes and neuronal activity-dependent polyadenylation site selection. *Neuron* 60:1022–1038. doi:10.1016/j.neuron.2008.11.029
55. Fortes P, Beloso A, Ortin J (1994) Influenza virus NS1 protein inhibits pre-mRNA splicing and blocks mRNA nucleocytoplasmic transport. *EMBO J* 13:704–712
56. Fox-Walsh K, Davis-Turak J, Zhou Y, Li H, Fu XD (2011) A multiplex RNA-seq strategy to profile poly(A+) RNA: application

- to analysis of transcription response and 3' end formation. *Genomics* 98:266–271. doi:10.1016/j.ygeno.2011.04.003
57. Fu Y, Dominissini D, Rechavi G, He C (2014) Gene expression regulation mediated through reversible m6A RNA methylation. *Nat Rev Genet* 15:293–306. doi:10.1038/nrg3724
 58. Fu Y, Sun Y, Li Y, Li J, Rao X, Chen C, Xu A (2011) Differential genome-wide profiling of tandem 3' UTRs among human breast cancer and normal cells by high-throughput sequencing. *Genome Res* 21:741–747. doi:10.1101/gr.115295.110
 59. Gehring NH, Frede U, Neu-Yilik G, Hundsdoerfer P, Vetter B, Hentze MW, Kulozik AE (2001) Increased efficiency of mRNA 3' end formation: a new genetic mechanism contributing to hereditary thrombophilia. *Nat Genet* 28:389–392
 60. Gilmartin GM (2005) Eukaryotic mRNA 3' processing: a common means to different ends. *Genes Dev* 19:2517–2521
 61. Graham RR, Kyogoku C, Sigurdsson S, Vlasova IA, Davies LRL, Baechler EC, Plenge RM, Koeuth T, Ortmann WA, Hom G, Bauer JW, Gillett C, Burt N, Cunninghame Graham DS, Onofrio R, Petri M, Gunnarsson I, Svenungsson E, Rönnblom L, Nordmark G, Gregersen PK, Moser K, Gaffney PM, Criswell LA, Vyse TJ, Syvänen A-C, Bohjanen PR, Daly MJ, Behrens TW, Altschuler D (2007) Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. *Proc Natl Acad Sci* 104:6758–6763. doi:10.1073/pnas.0701266104
 62. Graham DSC, Manku H, Wagner S, Reid J, Timms K, Gutin A, Lanchbury JS, Vyse TJ (2007) Association of IRF5 in UK SLE families identifies a variant involved in polyadenylation. *Hum Mol Genet* 16:579–591. doi:10.1093/hmg/dtl469
 63. Gruber AR, Martin G, Keller W, Zavolan M (2014) Means to an end: mechanisms of alternative polyadenylation of messenger RNA precursors. *Wiley Interdiscip Rev RNA* 5:183–196. doi:10.1002/wrna.1206
 64. Gruber AR, Martin G, Müller P, Schmidt A, Gruber AJ, Gumieny R, Mittal N, Jayachandran R, Pieters J, Keller W, van Nimwegen E, Zavolan M (2014) Global 3' UTR shortening has a limited effect on protein abundance in proliferating T cells. *Nat Commun* 5. doi:10.1038/ncomms6465
 65. Gunderson SI, Vagner S, Polycarpou-Schwarz M, Mattaj JW (1997) Involvement of the carboxyl terminus of vertebrate poly(A) polymerase in U1A autoregulation and in the coupling of splicing and polyadenylation. *Genes Dev* 11:761–773
 66. Hall-Pogar T, Liang S, Hague LK, Lutz CS (2007) Specific transacting proteins interact with auxiliary RNA polyadenylation elements in the COX-2 3'-UTR. *RNA* 13:1103–1115
 67. Hall-Pogar T, Zhang H, Tian B, Lutz CS (2005) Alternative polyadenylation of cyclooxygenase-2. *Nucleic Acids Res* 33:2565–2579
 68. Hampsey M, Reinberg D (2003) Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation. *Cell* 113:429–432
 69. Higgs DR, Goodbourn SE, Lamb J, Clegg JB, Weatherall DJ, Proudfoot NJ (1983) Alpha-thalassaemia caused by a polyadenylation signal mutation. *Nature* 306:398–400
 70. Hirose Y, Manley JL (2000) RNA polymerase II and the integration of nuclear events. *Genes Dev* 14:1415–1429
 71. Hogg JR, Goff SP (2010) Upfl senses 3'UTR length to potentiate mRNA decay. *Cell* 143:379–389. doi:10.1016/j.cell.2010.10.005
 72. Holler I, Grund K, Hentze MW, Kulozik AE (2014) MRNA 3' end processing: a tale of the tail reaches the clinic. *EMBO Mol Med* 6:16–26. doi:10.1002/emmm.201303300
 73. Hoque M, Ji Z, Zheng D, Luo W, Li W, You B, Park JY, Yehia G, Tian B (2013) Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. *Nat Methods* 10:133–139. doi:10.1038/nmeth.2288
 74. Hornyk C, Terzi LC, Simpson GG (2010) The spen family protein FPA controls alternative cleavage and polyadenylation of RNA. *Dev Cell* 18:203–213. doi:10.1016/j.devcel.2009.12.009
 75. Hu J, Lutz CS, Wilusz J, Tian B (2005) Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. *RNA* 11:1485–1493
 76. Jan CH, Friedman RC, Ruby JG, Bartel DP (2010) Formation, regulation and evolution of *Caenorhabditis elegans* 3'UTRs. *Nature*
 77. Jenal M, Elkon R, Loayza-Puch F, van Haaften G, Kuhn U, Menzies FM, Oude Vrielink JA, Bos AJ, Drost J, Rooijers K, Rubinsztein DC, Agami R (2012) The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* 149:538–553. doi:10.1016/j.cell.2012.03.022
 78. Ji Z, Lee JY, Pan Z, Jiang B, Tian B (2009) Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. *Proc Natl Acad Sci* 106:7028–7033. doi:10.1073/pnas.0900028106
 79. Ji Z, Tian B (2009) Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. *PLoS ONE* 4, e8419
 80. Johnson JM, Castle J, Garrett-Engele P, Kan Z, Loerch PM, Armour CD, Santos R, Schadt EE, Stoughton R, Shoemaker DD (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* 302:2141–2144
 81. Jonkers I, Lis JT (2015) Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev* 16:167–177. doi:10.1038/nrm3953
 82. Kahvejian A, Roy G, Sonenberg N (2001) The mRNA closed-loop model: the function of PABP and PABP-interacting proteins in mRNA translation. *Cold Spring Harb Symp Quant Biol* 66:293–300
 83. Kaida D, Berg MG, Younis I, Kasim M, Singh LN, Wan L, Dreyfuss G (2010) U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* 468:664–668. doi:10.1038/nature09479
 84. Kalsotra A, Cooper TA (2011) Functional consequences of developmentally regulated alternative splicing. *Nat Rev Genet* 12:715–729. doi:10.1038/nrg3052
 85. Kaneko S, Rozenblatt-Rosen O, Meyerson M, Manley JL (2007) The multifunctional protein p54^{nrb}/PSF recruits the exonuclease XRN2 to facilitate pre-mRNA 3' processing and transcription termination. *Genes Dev* 21:1779–1789
 86. Kaufmann I, Martin G, Friedlein A, Langen H, Keller W (2004) Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J* 23:616–626
 87. Ke S, Alemu EA, Mertens C, Gantman EC, Fak JJ, Mele A, Haribal B, Zucker-Scharff, Moore MJ, Park CY, Vagbo CB, Kussnierczyk A, Klungland A, Damell JE jr, Damell RB (2015) A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev* 1;29(19):2037–53
 88. Keller W, Minvielle-Sebastia L (1997) A comparison of mammalian and yeast pre-mRNA 3'-end processing. *Curr Opin Cell Biol* 9:329–336
 89. Kim YJ, Noguchi S, Hayashi YK, Tsukahara T, Shimizu T, Arahata K (2001) The product of an oculopharyngeal muscular dystrophy gene, poly(A)-binding protein 2, interacts with SKIP and stimulates muscle-specific gene expression. *Hum Mol Genet* 10:1129–1139
 90. Kleiman FE, Manley JL (1999) Functional interaction of BRCA1-associated BARD1 with polyadenylation factor CstF-50. *Science* 285:1576–1579
 91. Kuhn U, Wahle E (2004) Structure and function of poly(A) binding proteins. *Biochim Biophys Acta* 1678:67–84
 92. Kwan T, Benovoy D, Dias C, Gurd S, Provencher C, Beaulieu P, Hudson TJ, Sladek R, Majewski J (2008) Genome-wide analysis of transcript isoform variation in humans. *Nat Genet* 40:225–231

93. Kyburz A, Friedlein A, Langen H, Keller W (2006) Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3' end processing and splicing. *Mol Cell* 23:195–205
94. Lackford B, Yao CG, Charles GM, Weng LJ, Zheng XF, Choi EA, Xie XH, Wan J, Xing Y, Freudenberg JM, Yang PY, Jothi R, Hu G, Shi YS (2014) Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. *EMBO J* 33:878–889. doi:10.1002/embj.201386537
95. Lander ESea (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921
96. Langemeier J, Schrom EM, Rabner A, Radtke M, Zychlinski D, Saborowski A, Bohn G, Mandel-Gutfreund Y, Bodem J, Klein C, Bohne J (2012) A complex immunodeficiency is based on U1 snRNP-mediated poly(A) site suppression. *EMBO J* 31:4035–4044. doi:10.1038/emboj.2012.252
97. Lazarowitz SG, Compans RW, Choppin PW (1971) Influenza virus structural and nonstructural proteins in infected cells and their plasma membranes. *Virology* 46:830–843
98. Legendre M, Gautheret D (2003) Sequence determinants in human polyadenylation site selection. *BMC Genomics* 4:7
99. Lembo A, Di Cunto F, Provero P (2012) Shortening of 3'UTRs correlates with poor prognosis in breast and lung cancer. *PLoS ONE* 7, e31129. doi:10.1371/journal.pone.0031129
100. Li Y, Chen ZY, Wang W, Baker CC, Krug RM (2001) The 3'-end-processing factor CPSF is required for the splicing of single-intron pre-mRNAs in vivo. *RNA* 7:920–931
101. Li X, Manley JL (2006) Cotranscriptional processes and their influence on genome stability. *Genes Dev* 20:1838–1847
102. Lian Z, Karpikov A, Lian J, Mahajan MC, Hartman S, Gerstein M, Snyder M, Weissman SM (2008) A genomic analysis of RNA polymerase II modification and chromatin architecture related to 3' end RNA polyadenylation. *Genome Res* 18:1224–1237. doi:10.1101/gr.075804.107
103. Liao G-Y, An JJ, Gharami K, Waterhouse EG, Vanevski F, Jones KR, Xu B (2012) Dendritically targeted Bdnf mRNA is essential for energy balance and response to leptin. *Nat Med* 18:564–571. doi:10.1038/nm.2687
104. Lin Y, Li Z, Ozsolak F, Kim SW, Arango-Argoty G, Liu TT, Tenenbaum SA, Bailey T, Monaghan AP, Milos PM, John B (2012) An in-depth map of polyadenylation sites in cancer. *Nucleic Acids Res* 40:8460–8471. doi:10.1093/nar/gks637
105. Lou H, Gagel RF (1998) Alternative RNA processing—its role in regulating expression of calcitonin/calcitonin gene-related peptide. *J Endocrinol* 156:401–405
106. Lu Y, Qian XY, Krug RM (1994) The influenza virus NS1 protein: a novel inhibitor of pre-mRNA splicing. *Genes Dev* 8:1817–1828
107. Lukiw WJ, Bazan NG (1997) Cyclooxygenase 2 RNA message abundance, stability, and hypervariability in sporadic Alzheimer neocortex. *J Neurosci Res* 50:937–945
108. Lutz CS (2008) Alternative polyadenylation: a twist on mRNA 3' end formation. *ACS Chem Biol* 3:609–617. doi:10.1021/cb800138w
109. Lutz CS, Murthy KG, Schek N, O'Connor JP, Manley JL, Alwine JC (1996) Interaction between the U1 snRNP-A protein and the 160-kD subunit of cleavage-polyadenylation specificity factor increases polyadenylation efficiency in vitro. *Genes Dev* 10:325–337
110. MacDonald CC, McMahon KW (2010) Tissue-specific mechanisms of alternative polyadenylation: testis, brain, and beyond. *Wiley Interdiscip Rev RNA* 1:494–501. doi:10.1002/wrna.29
111. MacIsaac JL, Bogutz AB, Morrissy AS, Lefebvre L (2012) Tissue-specific alternative polyadenylation at the imprinted gene *Mest* regulates allelic usage at *Copg2*. *Nucleic Acids Res* 40:1523–1535. doi:10.1093/nar/gkr871
112. Mandel CR, Kaneko S, Zhang H, Gebauer D, Vethantham V, Manley JL, Tong L (2006) Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* 444:953–956
113. Mangone M, Manoharan AP, Thierry-Mieg D, Thierry-Mieg J, Han T, Mackowiak SD, Mis E, Zegar C, Gutwein MR, Khivansara V, Attie O, Chen K, Salehi-Ashtiani K, Vidal M, Harkins TT, Bouffard P, Suzuki Y, Sugano S, Kohara Y, Rajewsky N, Piano F, Gunsalus KC, Kim JK (2010) The landscape of *C. elegans* 3'UTRs. *Science* 329:432–435. doi:10.1126/science.1191244
114. Maniatis T, Reed R (2002) An extensive network of coupling among gene expression machines. *Nature* 416:499–506
115. Mapendano CK, Lykke-Andersen S, Kjems J, Bertrand E, Jensen TH (2010) Crosstalk between mRNA 3' end processing and transcription initiation. *Mol Cell* 40:410–422. doi:10.1016/j.molcel.2010.10.012
116. Martin G, Gruber Andreas R, Keller W, Zavolan M (2012) Genome-wide analysis of Pre-mRNA 3' end processing reveals a decisive role of human cleavage factor I in the regulation of 3' UTR length. *Cell Rep* 1:753–763. doi:10.1016/j.celrep.2012.05.003
117. Martincic K, Alkan SA, Cheatle A, Borghesi L, Milcarek C (2009) Transcription elongation factor ELL2 directs immunoglobulin secretion in plasma cells by stimulating altered RNA processing. *Nat Immunol* 10:1102–1109. doi:10.1038/ni.1786
118. Marzluft WF (2005) Metazoan replication-dependent histone mRNAs: a distinct set of RNA polymerase II transcripts. *Curr Opin Cell Biol* 17:274–280. doi:10.1016/j.ceb.2005.04.010
119. Masamha CP, Xia Z, Yang J, Albrecht TR, Li M, Shyu AB, Li W, Wagner EJ (2014) CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. *Nature* 510:412–416. doi:10.1038/nature13261
120. Matlin AJ, Clark F, Smith CWJ (2005) Understanding alternative splicing: towards a cellular code. *Nat Rev* 6:386–398
121. Mayr C, Bartel DP (2009) Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* 138:673–684
122. McCracken S, Longman D, Johnstone IL, Caceres JF, Blencowe BJ (2003) An evolutionarily conserved role for SRm160 in 3'-end processing that functions independently of exon junction complex formation. *J Biol Chem* 278:44153–44160
123. McLauchlan J, Phelan A, Loney C, Sandri-Goldin RM, Clements JB (1992) Herpes simplex virus IE63 acts at the posttranscriptional level to stimulate viral mRNA 3' processing. *J Virol* 66:6939–6945
124. McLauchlan J, Simpson S, Clements JB (1989) Herpes simplex virus induces a processing factor that stimulates poly(A) site usage. *Cell* 59:1093–1105
125. Millevoi S, Geraghty F, Idowu B, Tam JL, Antoniou M, Vagner S (2002) A novel function for the U2AF 65 splicing factor in promoting pre-mRNA 3'-end processing. *EMBO Rep* 3:869–874
126. Millevoi S, Loulergue C, Dettwiler S, Karaa SZ, Keller W, Antoniou M, Vagner S (2006) An interaction between U2AF 65 and CF I(m) links the splicing and 3' end processing machineries. *EMBO J* 25:4854–4864
127. Millevoi S, Vagner S (2010) Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic Acids Res* 38:2757–2774. doi:10.1093/nar/gkp1176
128. Moreira A, Takagaki Y, Brackenridge S, Wollerton M, Manley JL, Proudfoot NJ (1998) The upstream sequence element of the C2 complement poly(A) signal activates mRNA 3' end formation by two distinct mechanisms. *Genes Dev* 12:2522–2534
129. Moreira A, Wollerton M, Monks J, Proudfoot NJ (1995) Upstream sequence elements enhance poly(A) site efficiency of the C2

- complement gene and are phylogenetically conserved. *EMBO J* 14:3809–3819
130. Morris AR, Bos A, Diosdado B, Rooijers K, Elkon R, Bolijn AS, Carvalho B, Meijer GA, Agami R (2012) Alternative cleavage and polyadenylation during colorectal cancer development. *Clin Cancer Res* 18:5256–5266. doi:10.1158/1078-0432.ccr-12-0543
 131. Nagaike T, Logan C, Hotta I, Rozenblatt-Rosen O, Meyerson M, Manley JL (2011) Transcriptional activators enhance polyadenylation of mRNA precursors. *Mol Cell* 41:409–418. doi:10.1016/j.molcel.2011.01.022
 132. Natalizio BJ, Muniz LC, Arhin GK, Wilusz J, Lutz CS (2002) Upstream elements present in the 3'-untranslated region of collagen genes influence the processing efficiency of overlapping polyadenylation signals. *J Biol Chem* 277:42733–42740
 133. Nemeroff ME, Barabino SM, Li Y, Keller W, Krug RM (1998) Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3' end formation of cellular pre-mRNAs. *Mol Cell* 1:991–1000
 134. Niwa M, Rose SD, Berget SM (1990) In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes Dev* 4:1552–1559
 135. Nunes NM, Li W, Tian B, Furger A (2010) A functional human poly(A) site requires only a potent DSE and an A-rich upstream sequence. *EMBO J* 29:1523–1536
 136. Orkin SH, Cheng TC, Antonarakis SE, Kazazian HH Jr (1985) Thalassemia due to a mutation in the cleavage-polyadenylation signal of the human beta-globin gene. *EMBO J* 4:453–456
 137. Ozsolak F, Kapranov P, Foissac S, Kim SW, Fishilevich E, Monaghan AP, John B, Milos PM (2010) Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell* 143:1018–1029. doi:10.1016/j.cell.2010.11.020
 138. Phillips C, Jung S, Gunderson SI (2001) Regulation of nuclear poly(A) addition controls the expression of immunoglobulin M secretory mRNA. *EMBO J* 20:6443–6452
 139. Phillips C, Pachikara N, Gunderson SI (2004) U1A inhibits cleavage at the immunoglobulin M heavy-chain secretory poly(A) site by binding between the two downstream GU-rich regions. *Mol Cell Biol* 24:6162–6171
 140. Pinto PAB, Henriques T, Freitas MO, Martins T, Domingues RG, Wyrzykowska PS, Coelho PA, Carmo AM, Sunkel CE, Proudfoot NJ, Moreira A (2011) RNA polymerase II kinetics in polo polyadenylation signal selection. *EMBO J* 30:2431–2444. doi:10.1038/emboj.2011.156
 141. Poort SR, Rosendaal FR, Reitsma PH, Bertina RM (1996) A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood* 88:3698–3703
 142. Proudfoot NJ (2011) Ending the message: poly(A) signals then and now. *Genes Dev* 25:1770–1782. doi:10.1101/gad.17268411
 143. Proudfoot NJ, Furger A, Dye MJ (2002) Integrating mRNA processing with transcription. *Cell* 108:501–512
 144. Rao P, Yuan W, Krug RM (2003) Crucial role of CA cleavage sites in the cap-snatching mechanism for initiating viral mRNA synthesis. *EMBO J* 22:1188–1198
 145. Reed R, Hurt E (2002) A conserved mRNA export machinery coupled to pre-mRNA splicing. *Cell* 108:523–531
 146. Rehfeld A, Plass M, Krogh A, Friis-Hansen L (2013) Alterations in polyadenylation and its implications for endocrine disease. *Front Endocrinol* 4:53. doi:10.3389/fendo.2013.00053
 147. Rogers J, Early P, Carter C, Calame K, Bond M, Hood L, Wall R (1980) Two mRNAs with different 3' ends encode membrane-bound and secreted forms of immunoglobulin mu chain. *Cell* 20:303–312
 148. Rosonina E, Kaneko S, Manley JL (2006) Terminating the transcript: breaking up is hard to do. *Genes Dev* 20:1050–1056
 149. Ryan K, Calvo O, Manley JL (2004) Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. *RNA* 10:565–573
 150. Sachs AB, Sarnow P, Hentze MW (1997) Starting at the beginning, middle, and end: translation initiation in eukaryotes. *Cell* 89(6):831–838
 151. Sachs A, Wahle E (1993) Poly(A) tail metabolism and function in eucaryotes. *J Biol Chem* 268:22955–22958
 152. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB (2008) Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* 320:1643–1647. doi:10.1126/science.1155390
 153. Sandri-Goldin RM, Mendoza GE (1992) A herpesvirus regulatory protein appears to act post-transcriptionally by affecting mRNA processing. *Genes Dev* 6:848–863
 154. Schonemann L, Kuhn U, Martin G, Schafer P, Gruber AR, Keller W, Zavolan M, Wahle E (2014) Reconstitution of CPSF active in polyadenylation: recognition of the polyadenylation signal by WDR33. *Genes Dev* 28:2381–2393. doi:10.1101/gad.250985.114
 155. Schrijver I, Lenzi TJ, Jones CD, Lay MJ, Druzin ML, Zehnder JL (2003) Prothrombin gene variants in non-Caucasians with fetal loss and intrauterine growth retardation. *J Mol Diagn* 5:250–253
 156. Shatkin AJ, Manley JL (2000) The ends of the affair: capping and polyadenylation. *Nat Struct Biol* 7:838–842. doi:10.1038/79583
 157. Shell SA, Hesse C, Morris SM Jr, Milcarek C (2005) Elevated levels of the 64-kDa cleavage stimulatory factor (CstF-64) in lipopolysaccharide-stimulated macrophages influence gene expression and induce alternative poly(A) site selection. *J Biol Chem* 280:39950–39961
 158. Shepard PJ, Choi EA, Lu J, Flanagan LA, Hertel KJ, Shi Y (2011) Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq, RNA
 159. Shi Y (2012) Alternative polyadenylation: new insights from global analyses. *RNA* 18:2105–2117. doi:10.1261/rna.035899.112
 160. Shi Y, Di Giarmartino DC, Taylor D, Sarkeshik A, Rice WJ, Yates JR 3rd, Frank J, Manley JL (2009) Molecular architecture of the human pre-mRNA 3' processing complex. *Mol Cell* 33:365–376
 161. Shi Y, Manley JL (2015) The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. *Genes Dev* 29:889–897. doi:10.1101/gad.261974.115
 162. Singh P, Alley TL, Wright SM, Kamdar S, Schott W, Wilpan RY, Mills KD, Graber JH (2009) Global changes in processing of mRNA 3' untranslated regions characterize clinically distinct cancer subtypes. *Cancer Res* 69:9422–9430. doi:10.1158/0008-5472.can-09-2236
 163. Singh RK, Cooper TA (2012) Pre-mRNA splicing in disease and therapeutics. *Trends Mol Med* 18:472–482. doi:10.1016/j.molmed.2012.06.006
 164. Soetanto R, Hynes CJ, Patel H, Humphreys DT, Evers M, Duan GW, Parker BJ, Archer SK, Clancy JL, Graham RM, Beilharz TH, Smith NJ, Preiss T (2016) Role of miRNAs and alternative mRNA 3'-end cleavage and polyadenylation of their mRNA targets in cardiomyocyte hypertrophy. *Biochim Biophys Acta*. doi:10.1016/j.bbagn.2016.03.010
 165. Soo PY, Patel RK, Best S, Arya R, Thein SL (2005) Detection of prothrombin gene polymorphism at position 20209 (PT20209C/T): pilot study in a black population in the United Kingdom. *Thromb Haemost* 93:179–180
 166. Spies N, Burge CB, Bartel DP (2013) 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. *Genome Res* 23:2078–2090. doi:10.1101/gr.156919.113
 167. Spies N, Nielsen CB, Padgett RA, Burge CB (2009) Biased chromatin signatures around polyadenylation sites and exons. *Mol Cell* 36:245–254

168. Stacey SN, Sulem P, Jonasdottir A, Masson G, Gudmundsson J, Gudbjartsson DF, Magnusson OT, Gudjonsson SA, Sigurgeirsson B, Thorisdottir K, Ragnarsson R, Benediktsdottir KR, Nexo BA, Tjonneland A, Overvad K, Rudnai P, Gurzau E, Koppova K, Hemminki K, Corredera C, Fuentelsaz V, Grasa P, Navarrete S, Fuertes F, Garcia-Prats MD, Sanambrosio E, Panadero A, De Juan A, Garcia A, Rivera F, Planelles D, Soriano V, Requena C, Aben KK, van Rossum MM, Cremers RGHM, van Oort IM, van Spronsen D-J, Schalken JA, Peters WHM, Helfand BT, Donovan JL, Hamdy FC, Badescu D, Codreanu O, Jinga M, Csiki IE, Constantinescu V, Badea P, Mates IN, Dinu DE, Constantin A, Mates D, Kristjansdottir S, Agnarsson BA, Jonsson E, Barkardottir RB, Einarsson GV, Sigurdsson F, Moller PH, Stefansson T, Valdimarsson T, Johannsson OT, Sigurdsson H, Jonsson T, Jonasson JG, Tryggvadottir L, Rice T, Hansen HM, Xiao Y, Lachance DH, O'Neill BP, Kosel ML, Decker PA, Thorleifsson G, Johannsdottir H, Helgadottir HT, Sigurdsson A, Steinthorsdottir V, Lindblom A, Sandler RS, Keku TO, Banasik K, Jorgensen T, Witte DR, Hansen T, Pedersen O, Jinga V, Neal DE, Catalona WJ, Wrensch M, Wiencke J, Jenkins RB, Nagore E, Vogel U, Kiemeny LA, Kumar R, Mayordomo JI, Olafsson JH, Kong A, Thorsteinsdottir U, Rafnar T, Stefansson K (2011) A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. *Nat Genet* 43:1098–1103. <http://www.nature.com/ng/journal/v43/n11/abs/ng.926.html-supplementary-information>
169. Stutz F, Rosbash M (1998) Nuclear RNA export. *Genes Dev* 12:3303–3319
170. Takagaki Y, Manley JL (1998) Levels of polyadenylation factor CstF-64 control IgM heavy chain mRNA accumulation and other events associated with B cell differentiation. *Mol Cell* 2:761–771
171. Takagaki Y, Seipelt RL, Peterson ML, Manley JL (1996) The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell* 87:941–952
172. Thomas LF, Saetrom P (2012) Single nucleotide polymorphisms can create alternative polyadenylation signals and affect gene expression through loss of microRNA-regulation. *PLoS Comput Biol* 8, e1002621. doi:10.1371/journal.pcbi.1002621
173. Tian B, Hu J, Zhang H, Lutz CS (2005) A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res* 33:201–212
174. Tian B, Manley JL (2013) Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem Sci* 38:312–320. doi:10.1016/j.tibs.2013.03.005
175. Vagner S, Vagner C, Mattaj JW (2000) The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing. *Genes Dev* 14:403–413
176. Venkataraman K, Brown KM, Gilmartin GM (2005) Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes Dev* 19:1315–1327
177. Veraldi KL, Arhin GK, Martincic K, Chung-Ganster LH, Wilusz J, Milcarek C (2001) hnRNP F influences binding of a 64-kilodalton subunit of cleavage stimulation factor to mRNA precursors in mouse B cells. *Mol Cell Biol* 21:1228–1238
178. Vickers TA, Crooke ST (2012) siRNAs targeted to certain polyadenylation sites promote specific, RISC-independent degradation of messenger RNAs. *Nucleic Acids Res* 40:6223–6234. doi:10.1093/nar/gks239
179. Vorlova S, Rocco G, Lefave CV, Jodelka FM, Hess K, Hastings ML, Henke E, Cartegni L (2011) Induction of antagonistic soluble decoy receptor tyrosine kinases by intronic polyA activation. *Mol Cell* 43:927–939. doi:10.1016/j.molcel.2011.08.009
180. Wahle E, Kuhn U (1997) The mechanism of 3' cleavage and polyadenylation of eukaryotic pre-mRNA. *Prog Nucleic Acid Res Mol Biol* 57:41–71
181. Wang G-S, Cooper TA (2007) Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet* 8:749–761
182. Wang L, Dowell RD, Yi R (2013) Genome-wide maps of polyadenylation reveal dynamic mRNA 3'-end formation in mammalian cell lineages. *RNA* 19:413–425. doi:10.1261/ma.035360.112
183. Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* 456:470–476
184. Wassaman KM, Steitz JA (1993) Association with terminal exons in pre-mRNAs: a new role for the U1 snRNP? *Genes Dev* 7:647–659
185. Weiner AM (2005) E Pluribus Unum: 3' end formation of polyadenylated mRNAs, histone mRNAs, and U snRNAs. *Mol Cell* 20:168–170
186. Wiestner A, Tehrani M, Chiorazzi M, Wright G, Gibellini F, Nakayama K, Liu H, Rosenwald A, Muller-Hermelink HK, Ott G, Chan WC, Greiner TC, Weisenburger DD, Vose J, Armitage JO, Gascoyne RD, Connors JM, Campo E, Montserrat E, Bosch F, Smeland EB, Kvaloy S, Holte H, Delabie J, Fisher RI, Grogan TM, Miller TP, Wilson WH, Jaffe ES, Staudt LM (2007) Point mutations and genomic deletions in CCND1 create stable truncated cyclin D1 mRNAs that are associated with increased proliferation rate and shorter survival. *Blood* 109:4599–4606. doi:10.1182/blood-2006-08-039859
187. Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, Penkett CJ, Rogers J, Bahler J (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* 453:1239–1243
188. Wilkening S, Pelechano V, Jarvelin AI, Tekkedil MM, Anders S, Benes V, Steinmetz LM (2013) An efficient method for genome-wide polyadenylation site mapping and RNA quantification. *Nucleic Acids Res* 41, e65. doi:10.1093/nar/gks1249
189. Wilusz JE, Spector DL (2010) An unexpected ending: noncanonical 3' end processing mechanisms. *RNA* 16:259–266. doi:10.1261/ma.1907510
190. Wood AJ, Schulz R, Woodfine K, Koltowska K, Beechey CV, Peters J, Bourc'his D, Oakey RJ (2008) Regulation of alternative polyadenylation by genomic imprinting. *Genes Dev* 22:1141–1146. doi:10.1101/gad.473408
191. Xia Z, Donehower LA, Cooper TA, Neilson JR, Wheeler DA, Wagner EJ, Li W (2014) Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types. *Nat Commun* 5. doi:10.1038/ncomms6274
192. Xiang K, Tong L, Manley JL (2014) Delineating the structural blueprint of the pre-mRNA 3'-end processing machinery. *Mol Cell Biol* 34:1894–1910. doi:10.1128/MCB.00084-14
193. Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, Lander ES, Kellis M (2005) Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* 434:338–345
194. Yan J, Marr TG (2005) Computational analysis of 3'-ends of ESTs shows four classes of alternative polyadenylation in human, mouse, and rat. *Genome Res* 15:369–375
195. Yao P, Potdar Alka A, Arif A, Ray Partho S, Mukhopadhyay R, Willard B, Xu Y, Yan J, Saidel Gerald M, Fox Paul L (2012) Coding region polyadenylation generates a truncated tRNA synthetase that counters translation repression. *Cell* 149:88–100. doi:10.1016/j.cell.2012.02.018
196. Yoon OK, Hsu TY, Im JH, Brem RB (2012) Genetics and regulatory impact of alternative polyadenylation in human B-lymphoblastoid cells. *PLoS Genet* 8, e1002882. doi:10.1371/journal.pgen.1002882
197. Zhang H, Lee JY, Tian B (2005) Biased alternative polyadenylation in human tissues. *Genome Biol* 6:R100

198. Zhao J, Hyman L, Moore C (1999) Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev* 63: 405–445