Title: Alternative polyadenylation in the regulation and dysregulation of gene expression

Authors: Rachael Emily Turner, Andrew David Pattison, Traude Helene Beilharz

PII: S1084-9521(17)30340-3
DOI: http://dx.doi.org/10.1016/j.semcdb.2017.08.056
Reference: YSCDB 2365

To appear in: *Seminars in Cell & Developmental Biology*

Received date: 14-6-2017
Revised date: 30-8-2017
Accepted date: 30-8-2017

Please cite this article as: Turner Rachael Emily, Pattison Andrew David, Beilharz Traude Helene.Alternative polyadenylation in the regulation and dysregulation of gene expression.*Seminars in Cell and Developmental Biology* http://dx.doi.org/10.1016/j.semcdb.2017.08.056

Alternative polyadenylation in the regulation and dysregulation of gene expression.

Rachael Emily Turner*, Andrew David Pattison* and Traude Helene Beilharz.

* Denotes equal author contribution

Development and stem cells Program, Monash Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Victoria Australia, 3800.

Correspondence should be addressed to Traude Beilharz. Tel: +613 9902 9183; Fax: +613 39902 9500; Email: traude.beilharz@monash.edu

Introduction

Advances in RNA analysis technologies have led to a new level of appreciation for the complexity of mRNA metabolism. In the area of 3'-end processing, increasing sophistication of first tiling arrays [1-3] and then in next generation sequencing approaches [4-11] has revealed that alternative polyadenylation (APA) is common in eukaryotic mRNA. APA refers to situations where more than one potential polyadenylation site exists within the 3'UTR of a mRNA molecule. The result of this 3'-end diversity, is a highly expanded regulatory and protein coding repertoire over what was previously thought. This provides scope for sophisticated regulatory paradigms in normal growth and development but also means relatively minor changes to mRNA processing can have major implications in disease. Here we will discuss the mechanisms of alternative polyadenylation and how its dysregulation can lead to disease.

RNA Processing for Alternative Polyadenylation

Where a nascent mRNA is associated with a single encoded cleavage and polyadenylation site it is termed constitutive polyadenylation (Figure 1a). Such mRNA polyadenylation can be further classified into four distinct groups based on the position of the poly(A) sites (Figure 1b-e). The most frequent form of APA is tandem 3'UTR APA, where multiple mRNA isoforms are produced that vary only in the length of their 3'UTR. This version is generally thought to be associated with altered stability, translational efficiency and/or localisation between 3'UTR mRNA isoforms. Unlike tandem 3'UTR APA, the three other forms involve changes to the protein-coding potential and therefore exist under the umbrella term, coding site (CDS) APA. In this case, APA events are linked to alternative splicing and result in distinct protein products between isoforms [12, 13]. Firstly, terminal exon APA involves the incorporation of alternative terminal exons into the nascent transcript through the use of different poly(A) sites between isoforms. Secondly, alternative intronic APA results from the recognition of cryptic cleavage sites present within introns. This transforms

the normally non-coding region upstream of this site into a composite exon [12]. Lastly, exons other than the terminal exon can also contain cryptic poly(A) sites and this is referred to as internal exon APA. Unlike alternative terminal exon APA, mRNA transcripts produced from internal exon APA will not possess a stop codon unless the site of cleavage directly follows a T or TG residue that is converted into a stop-codon by the addition of a poly(A)-tract. In the absence of a stop, no 3'UTR is generated, and since A-tracts generate C-terminal poly-lysine tags, any resulting protein products are likely targeted for decay [14]. Intronic APA can also suffer this fate, in the absence of an alternative stop codon within the retained part of the intron, upstream of the poly(A) site. It is possible for more than one type of APA to exist for a single transcript.

Mechanisms of APA Choice: First Come, First Served or Survival of the Fittest?

Despite the pervasive nature of APA, how any particular site is specified for cleavage under a given circumstance is still far from clear. Albeit, several models have emerged. Firstly, it has been proposed that the proximal cleavage site has an intrinsic temporal advantage over the distal site [15]. Early work by Denome and Cole [16] demonstrated that when two identical polyadenylation signals were present in the 3'UTR of a reporter construct, increasing the distance between these sites biased poly(A)-site usage further toward proximal by increasing its temporal advantage. Further, in the ‘enhancer of rudimentary’ e(r) transcript in Drosophila, the proximal cleavage site is preferentially used in males, whereas the distal site is utilised in females. Switching the sequence of these sites did not alter processing [17]. This suggests that position can trump sequence, perhaps simply because a proximal site is transcribed first and therefore has more time to be recognised by the cleavage and polyadenylation machinery. This idea has become the “first come, first served” model [18].

This proximal advantage can be influenced by the transcriptional elongation rate. When APA was monitored in an RNA polymerase II mutant with a slower elongation rate, an aberrant and functionally deleterious proximal cleavage site in the Polo transcript of *D. melanogaster* was preferentially used [19]. This increased proximal usage was interpreted as allowing more time for the proximal site to be recognised by the 3'end processing machinery. If poly(A)-site choice then depends on transcriptional elongation rate, it does not appear to generally correspond to a gene's transcriptional frequency. In both the human and mouse transcriptome, more abundant genes tend to have shorter 3'UTRs than their more lowly expressed counterparts [20], a finding that was recapitulated in reporter assays where expression from stronger promoters caused a preference for proximal site choice [20]. An exception to this general rule is seen in yeast where strong transcriptional up-regulation can shift cleavage to a distal site [21]. This might be explained by the combined effect of very short 3'UTRs in yeast and a high transcriptional rate, such that the 3'-end machinery 'misses the moment' for cleavage at the proximal site, cleaving at a distal site once it gets the chance. Given that the cleavage and adenylation machinery is thought to travel with the polymerase, perhaps it is only once transcription slows, that end formation can occur.

RNA polymerase II pausing has been correlated with switches in cleavage-site choice. Pausing at the poly(A) signal was shown to be heightened at more highly expressed genes indicating that even a local decrease in elongation rate increases proximal site usage [20]. In a further example of this kinetic model for poly(A)-choice, deletion of a pause site downstream of the proximal cleavage site in the IgM gene causes increased use of the more distal site [22]. Such stalls can come from structured DNA or RNA elements. Moreover, local chromatin structure and epigenetic marks impact alternative polyadenylation. Poly(A)-sites are associated with a strong depletion of nucleosomes whereas downstream regions are enriched for nucleosomes [23-25]. For genes with multiple sites, the more highly used site was associated with a lower nucleosome occupancy directly surrounding the poly(A) site and more pronounced nucleosome enrichment downstream

[25]. This may be due to altered RNA polymerase II kinetics or to the presence of a more favourable environment for efficient assembly of the cleavage and polyadenylation machinery [26]. Finally, the fact that epigenetics can play a role is demonstrated in findings that cleavage site choice can occur in an allele-specific manner. For the mouse imprinted gene H13, the alternative poly(A)-sites are separated by a CpG island. However, this is only methylated on the maternal copy. This methylation causes stalling of RNA polymerase II at this site only on the maternal allele, biasing cleavage toward the proximal site, whereas distal sites are selected in the absence of methylation at the paternal allele [27]. Such findings have prompted an alternative model, being that of 'survival of the fittest'.

The "survival of the fittest" model explains changes in cleavage site choice that favour the use of more distal poly(A) sites. This model focuses on findings showing that the positioning and efficiency of 3'-end cleavage and polyadenylation are largely determined by the interaction of cleavage and polyadenylation factors with *cis* elements surrounding the potential poly(A)-site. Several core *cis* elements within the pre-mRNA appear to be important for 3'-end processing. Proudfoot and Brownlee [28] discovered a conserved AAUAAA hexamer in the region upstream of the cleavage site in metazoans (Figure 2). This element has been termed the polyadenylation signal (PAS) and typically occurs within 40 nucleotides of the cleavage site [29]. Although AAUAAA is the canonical PAS sequence, variants of this hexamer are also observed in metazoans [5, 30, 31]. In general, deviations from the canonical sequence are associated with weaker PAS [reviewed: [32]]. Such variant PAS are more frequent in genes with multiple polyadenylation sites [33, 34] with the canonical PAS tending to occur at distal sites, and proximal sites often being variant signals [5, 35].

In addition to the PAS element, there are distinct U-rich and GU-rich elements located within 100 nucleotides downstream (DSE) of the cleavage site [36]. An upstream U-rich element (USE) containing UGUA motifs also tends to be positioned within 40 to 100 nucleotides of the cleavage

site [37]. This *cis* element pattern is generally conserved, however, higher variation in signals are seen for plants and the budding yeast *Saccharomyces cerevisiae* [38]. Together, these core motifs along with auxiliary elements help to define the cleavage site of the mRNA and impact the efficiency of 3'end processing. It has been suggested that these elements act cooperatively so that the absence of, or weak sequences for one element may be compensated for by stronger elements also present at that cleavage site [38]. To this end, sites lacking a PAS element tend to rely on strong USEs, DSEs or auxiliary elements to bind the 3'end processing machinery [33, 39].

In general, the more distal poly(A) sites utilise stronger *cis* elements than proximal sites and are also more likely to be conserved [35]. Thus, proximal sites tend to be weaker than distal sites and are less likely to be recognised by the 3'end processing machinery [40]. However favourable the sequence configuration in *cis*, the actual site that is chosen for cleavage and polyadenylation is strongly impacted by the concentration of the core cleavage and polyadenylation factors as well as other proteins.

The Influence of Cleavage Factor Concentration on Poly(A)-site choice

A complex machinery of protein factors is involved in the 3'-end processing of eukaryotic pre-mRNA. This consists of 20 characterised proteins in yeast [41] and perhaps more than 80 proteins in human cells [42]. The core factors involved in cleavage and polyadenylation in these two organisms are generally conserved, although there are some differences in the consensus sequence for PAS recognition and sub-complex organisation [43] (See also Table 1). The mammalian 3'-end processing machinery is comprised of four core sub-complexes. These include cleavage and polyadenylation specification factor (CSPF), cleavage stimulation factor (CstF), cleavage factor I (CFIm) and cleavage factor II (CFIIm) [44]. These factors bind selectively to the cleavage and polyadenylation site prior to any reaction taking place [43]. The CPSF, CstF and CFIm complexes

recognise and bind to the PAS, DSE and USE respectively [45]. These three factors then recruit CFIIm and other proteins such as the polyadenylation polymerase (Pap), symplekin, and polyadenylated-binding nuclear protein 1 (PABPN1) to form the 3'-end processing machinery and allow cleavage and polyadenylation to take place [46]. In contrast to the mammalian system, the yeast machinery is comprised of three main complexes including cleavage and polyadenylation factor (CPF), cleavage factor IA (CFIA) and cleavage factor IB (CFIB) [47]. CFIA contains homologous subunits to those in mammalian CFIIm and CstF, except for the apparent absence of a CstF-50 equivalent [47]. Similarly, CPF contains subunits that are homologous to those in mammalian CPSF. However, these are distributed into the two different sub-complexes, CFII and PFI, which make up CPF [47]. Some factors also appear to be unique to either mammals or yeast. The yeast CFIB factor, for example, appears to lack a homologue in mammals whereas the mammalian factor CFIm appears to be absent from yeast [48]. The core cleavage factors in both mammalian and yeast systems are indicated in Table 1.

The effect of this machinery's concentration on the regulation of APA was first reported by Takagaki et al. [49, 50]. They demonstrated that in resting B cells, the low levels of CstF-64 were associated with the preferential use of the distal cleavage site in IgM mRNAs, which, contained a strong CstF-64 binding site. However, during B cell activation, elevated levels of CstF-64 correlated with a switch to the proximal site with weak CstF-64 binding capability. This suggests that higher levels of this 3'-end processing factor promote the recognition of weaker cleavage sites whilst limiting concentrations cause the preferential use of the stronger sites. This has since been shown to be a global trend for CstF-64 [40, 51]. However, the overall effect of CstF-64 knockdown is relatively small compared to the co-depletion of CstF-64 and its variant form CstF-64τ [40]. As CstF-64τ depletion alone also has a comparatively mild impact on cleavage site switching, this suggests that CstF-64 and CstF-64τ have at least partially redundant functions [52]. Similarly, reduced levels of the CPSF factor Fip1 or yeast Hrp1 have been associated with increased distal site

usage [53, 54]. In striking contrast to this, depletion of the CFIm subunits CFIm-25 or CFIm-68 causes an increased use of more proximal cleavage sites [29, 55-58]. This implies that unlike the other core 3'-end processing complexes, CFIm normally represses cleavage at the proximal site. The cellular balance of 3'-end processing sub-machineries is therefore of critical importance to APA choice.

Together these data indicate that the strength of mRNA *cis* elements and the concentration of 3'-end processing factors together define the pattern of APA for many genes. However, it is of note that when target mRNA affected by changes in cleavage machinery concentrations were compared, there was little overlap, suggesting that 3'-end processing factors each influence the alternative polyadenylation of a specific subset of genes [53] (see also Figure 2). Another possibility is that each different experimental system that has identified shifts in 3'UTR choice [29, 40, 59] is prone to unique gene regulatory paradigms. A clearer understanding will require systematic depletion of the machinery within a single experimental system [13], and to understand the native 3'-end profile differences between systems.

Coupling Between Transcriptional and 3'-end Processing Machineries

Functional coupling between transcription and 3'-end processing extends beyond the role of the transcriptional elongation rate in cleavage site choice. RNA polymerase II is an essential mRNA polyadenylation factor in its own right, both *in vivo* and *in vitro* [60, 61]. The RNA polymerase II carboxyl-terminal domain (CTD) interacts with many of the 3'-end processing factors [62-65]. Interaction with the CTD is thought to recruit the 3'-end processing sub-complexes to the pre-mRNA, positioning them for cleavage and polyadenylation, with RNA polymerase II acting as a platform for 3'-end processing. Furthermore, various transcription factors have also been shown to play a role in cleavage site selection. PAF1C is a transcriptional elongation factor that has been

associated with enhanced 3'-end processing [66]. Depletion of the PAF1C subunits Cdc73, Paf1 or Ski8 resulted in a global increase in proximal site usage in murine myocytes [67]. Paf1 ablation was also linked to the accumulation of RNA polymerase II along gene bodies suggesting that this is due to a reduced transcriptional elongation rate [67]. Similarly, decreased expression of the murine transcription elongation factor Ell2 also caused enhanced proximal site usage and a concomitant switch from membrane-tethered immunoglobulins to their secreted form [68]. In this case, Ell2 increased CstF-64 association with RNA polymerase II and thus caused a high local concentration of CstF-64 at more proximal cleavage sites.

Coupling Between Splicing and 3'-end Processing Machineries

A link between alternative polyadenylation and splicing is becoming increasingly clear (Figure 1c-e). For example, the U1 snRNP affects cleavage and polyadenylation independently of its role in splicing. Knockdown of U1 snRNP promotes the use of cryptic polyadenylation sites within introns close to the 5'-end of the transcript [69]. It was suggested that binding of U1 snRNP to these regions blocks their recognition. Surprisingly, more moderate decreases in U1 snRNP levels elicited a shift to more proximal 3'UTR cleavage sites rather than the use of upstream intronic cleavage sites indicating that the impact of U1 snRNP is dose-dependent [70]. Furthermore, U2 snRNP mediates CPSF loading to the pre-mRNA and the U2 snRNP auxiliary factor (U2AF) interacts with CFIm-59, both of which stimulate cleavage and polyadenylation [71, 72]. A confounding factor to the interpretation of links between splicing and polyadenylation is the degree of cross-talk between the two systems [73]. Movassat et al. [74], suggest for example, that extensive terminal exon splicing induced by CstF-64 knockdown can be explained by changes to 3'UTR choice in core splicing factors, the altered expression of which then indirectly alter transcriptome-wide splicing choices.

Regulation of Alternative Polyadenylation by PAS Occlusion

Polyadenylation sites can be blocked by protein and/or RNA elements that compete with the binding of 3'-end processing factors. Multiple proteins have been identified that interfere with CstF-64 binding and cause a shift in polyadenylation site choice [73, 75, 76]. For example, the poly(A) binding protein nuclear 1 (PABPN1) plays an important role in 3'end processing [77]. In addition to its function in poly(A)-length control, PABPN1 has been shown to associate with weaker proximal APA sites suppressing their cleavage [59]. As a result, PABPN1 knockdown facilitates a shift to proximal site usage. Depletion of the major cytoplasmic poly(A) binding protein, PABPC1, appears to cause 3'UTR shortening to a similar extent as PABPN1 indicating that this may be a general function of PABPs [13]. Conversely, some RNA binding proteins aid the 3'-end processing machinery in site recognition, promoting the use of weaker *cis* elements. The human TREX subunit Thoc5 co-transcriptionally loads CFIm-68 onto target genes [78]. Therefore, like CFIm-68 depletion, a decrease in Thoc5 results in a shift towards proximal site usage [78]. Furthermore, the cytoplasmic polyadenylation element binding protein 1 (CPEB1) can shuttle to the nucleus and bind to cytoplasmic polyadenylation elements (CPE) within the pre-mRNA sequence [79]. This aids in the recruitment of CPSF to weaker DSEs and therefore proximal cleavage site use is increased. Importantly, occluded cleavage sites can be responsive to, and revealed by, cellular signalling. Danckwardt *et al*. [80], identified the RNA binding proteins FBP2 and FBP3, at upstream elements within the F2 (thrombin) 3'UTR. External stresses such as inflammatory cytokines that activate p38 MAPK, result in phosphorylation of the FBP RNA binding proteins, causing their dissociation from target mRNA, and activation of 3'-end processing [80].

An emerging theme in modulators of APA is to control their own cleavage and polyadenylation in an auto-regulatory fashion. For example, the Drosophila embryonic lethal abnormal visual system protein (ELAV) and its homologue in mammals, HuR, directly bind to proximal cleavage sites

suppressing 3'end processing through steric hindrance [81, 82]. Both ELAV and HuR auto-regulate their abundance through binding to proximal polyadenylation sites within their own pre-mRNAs causing longer 3'UTR isoforms [83, 84]. The cleavage and polyadenylation machinery itself also appears to be self-equilibrating. The CstF-77 primary transcript for example, harbours a cryptic truncating adenylation site that buffers expression of the full-length, functional protein [85]. Thus an excess of cleavage and polyadenylation activity feeds back to limit its own expression.

Pre-mRNA secondary structures have also been shown to impact site selection [86, 87]. IEAAT2 possesses a stem-loop structure within its 7^{th} intron. Adenosine/Inosine RNA editing results in increased use of a cryptic alternative polyadenylation site within the stem-loop region [88]. Finally, non-coding RNAs may also play a role in polyadenylation site switching. The long non-coding RNA colon cancer-associated transcript 2 (CCAT2), has been implicated in cleavage site selection for the gene glutaminase (GLS) [89]. CCAT2 interacts with the CFIm complex, causing CFIm-25 recruitment, which, in this case, promotes cleavage at a cryptic PAS within intron 14 of GLS. This results in the synthesis of the short GAC protein isoform that triggers glutamine metabolism and is implicated in metastasis.

Rather than any of these models acting alone, it is likely that alternative polyadenylation site selection is under the combinatorial control of various mechanisms and that site choice can be regulated in a multitude of ways including those yet to be identified.

Dysregulation of Alternative Polyadenylation in Disease

As the majority of human genes undergo alternative 3'-end processing at some time during development [7], it is not surprising that APA dysregulation has been associated with multiple disease states (See Table 2). In general, genome-wide research in disease tissue and in models of

disease have suggested a trend toward global shortening of 3'UTRs. Although this trend is being challenged [90]. Furthermore, it is important to note that comparison of the specific transcripts subject to change in 3'-end usage in these studies shows surprisingly little overlap. Of the 3137 different transcripts identified subject to APA across 4 different studies [51, 90-92] compared here, only 10 are shared, and of these only 7 change in the same direction, toward shortening (Figure 3a). Since the discovery that proliferating and cancer cells have a tendency to express shorter 3'UTRs [93, 94], the role of APA in cancer has become the most well-studied of any disease state [1, 58, 95]. In a comprehensive recent review, Gruber et al. [96] provide evidence that the expression level of 3'-end processing factors is highly correlated to the proliferative index of human cells. Given that high levels of such factors tend to alter APA site selection, it is not surprising that APA in cancer has been a major topic of current research.

Perhaps the most comprehensive study of APA in cancer to date was performed by Xia et al. [51], who inferred APA events from the RNA-Seq read coverage of tumour/normal pairs in the cancer genome atlas (TCGA). Their bioinformatic analysis identified 1,346 dynamic APA events across 7 tumour types. In general, 3'UTR shortening was found to be associated with tumourigenesis and CstF-64 was suggested to be a master regulator of APA in tumours. The authors also found APA to be more predictive of tumour outcome than gene expression. The possibility of the prognostic value of APA events was independently shown also in prostate cancer by Li et al. [95], albeit, these authors also note a significant sub-population of transcripts that switch to longer 3'UTR isoforms. The heterogeneity of APA events is underpinned by comparison of the overlap of APA events in different cancer types from the study by Xia et al. [51] (Figure 3b). Of the 920 transcripts subject to APA, only 22 are shared between the five cancers shown. Of these 22, 18 are shortened in all cancer types.

Cancer can be a heterogeneous disease. In breast cancer, for example, multiple subtypes are known, with varying properties determined from both clinical markers [97] and associated gene expression signatures [98]. It remains to be seen whether APA could provide additional prognostic value over current clinical parameters in breast cancer. To date, the research has been mixed. By contrasting two different breast cancer cell lines to a cultured mammary epithelial cell line MCF10A, Fu et al. [92], showed a polar opposite trend in 3'UTR choices. The luminal derived MCF7 cell line exhibited the expected 3'UTR shortening. In contrast, the MDA-MB-231 cells showed lengthening relative to the epithelial control. Such data highlight the complexity of the underlying biology in each individual tumour type and suggest that definitive global statements are no longer useful as descriptors in disease. Instead, new research should focus on identification of the specific APA events that are prognostic for disease outcome.

Importantly, as with alternative splicing, alternative 3'-end processing can have major regulatory effects with very little change in overall mRNA expression level [99]. The addition/removal, of stability, translation and localisation elements in just a few key regulatory genes could send gene expression programs down divergent developmental trajectories. If this were not already complicated enough, new research shows that 3'UTRs can act as protein complex assembly scaffolds [100, 101]. In this case, a change in 3'-end usage can impact the co-translational assembly of cellular machines. It is therefore important not to be drawn into generalisations. Although some 3'UTR shortening events clearly result in increases in protein expression (See Table 2), multiple studies now show that shortened 3'UTRs do not necessarily result in more protein synthesis [90]. Moreover, some research suggests that distal isoforms can be more efficiently translated [19, 102, 103]. Indeed there may be as many activating elements in 3'UTRs as there are repressive elements [104].

While the number of protein-coding genes are maintained, APA events appear to scale with complexity, [Reviewed: [100]], so it follows that APA is likely a means of increasing functional and regulatory diversity. The dysregulation of these processes is therefore likely common to all complex diseases. Systematic identification of the highly diversified and specific disease-associated APA events provides a massive challenge. However, if the early suggestions of the prognostic power of specific APA events in cancer prove correct, there is also a considerable opportunity. This might be particularly the case when APA events are considered in combination with other modes of gene expression control such as alternative splicing and post transcriptional regulatory mechanisms. Finally, in analogy to the breakthroughs in exon-skipping for muscular dystrophy, perhaps future therapeutics that selectively switch 3'UTR choice, will provide an opportunity to produce among the most targeted, medicines ever created.

Acknowledgements

We thank members of the Beilharz Lab for constructive comment on the manuscript. RET was supported by an Australian Postgraduate scholarship, ADP was supported by a Monash Postgraduate scholarship and THB was supported by a Monash Biodiscovery Fellowship. Research in the Beilharz Lab was funded by grants from the NHMRC (APP1128250) and ARC (DP170100569).

References

[1] A. Lembo, F. Di Cunto, P. Provero, Shortening of 3'UTRs correlates with poor prognosis in breast and lung cancer, PLoS One 7(2) (2012) e31129.
[2] Z. Ji, B. Tian, Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types, PLoS One 4(12) (2009) e8419.

[3] J.Y. Park, W. Li, D. Zheng, P. Zhai, Y. Zhao, T. Matsuda, S.F. Vatner, J. Sadoshima, B. Tian, Comparative analysis of mRNA isoform expression in cardiac hypertrophy and development reveals multiple post-transcriptional regulatory modules, PLoS One 6(7) (2011) e22391.
[4] O.K. Yoon, R.B. Brem, Noncanonical transcript forms in yeast and their regulation during environmental stress, RNA 16(6) (2010) 1256-67.
[5] P.J. Shepard, E.A. Choi, J. Lu, L.A. Flanagan, K.J. Hertel, Y. Shi, Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq, RNA 17(4) (2011) 761-72.
[6] A.H. Beck, Z. Weng, D.M. Witten, S. Zhu, J.W. Foley, P. Lacroute, C.L. Smith, R. Tibshirani, M. van de Rijn, A. Sidow, R.B. West, 3'-end sequencing for expression quantification (3SEQ) from archival tumor samples, PLoS One 5(1) (2010) e8768.
[7] A. Derti, P. Garrett-Engele, K.D. Macisaac, R.C. Stevens, S. Sriram, R. Chen, C.A. Rohl, J.M. Johnson, T. Babak, A quantitative atlas of polyadenylation in five mammals, Genome Res 22(6) (2012) 1173-83.
[8] F. Ozsolak, P. Kapranov, S. Foissac, S.W. Kim, E. Fishilevich, A.P. Monaghan, B. John, P.M. Milos, Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation, Cell 143(6) (2010) 1018-29.
[9] S. Wilkening, V. Pelechano, A.I. Jarvelin, M.M. Tekkedil, S. Anders, V. Benes, L.M. Steinmetz, An efficient method for genome-wide polyadenylation site mapping and RNA quantification, Nucleic Acids Res 41(5) (2013) e65.
[10] M. Hoque, Z. Ji, D. Zheng, W. Luo, W. Li, B. You, J.Y. Park, G. Yehia, B. Tian, Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing, Nat Methods 10(2) (2013) 133-9.
[11] P.F. Harrison, D.R. Powell, J.L. Clancy, T. Preiss, P.R. Boag, A. Traven, T. Seemann, T.H. Beilharz, PAT-seq: a method to study the integration of 3'-UTR dynamics with gene expression in the eukaryotic transcriptome, RNA 21(8) (2015) 1502-10.
[12] B. Tian, Z. Pan, J.Y. Lee, Widespread mRNA polyadenylation events in introns indicate dynamic interplay between polyadenylation and splicing, Genome Res 17(2) (2007) 156-65.
[13] W. Li, B. You, M. Hoque, D. Zheng, W. Luo, Z. Ji, J.Y. Park, S.I. Gunderson, A. Kalsotra, J.L. Manley, B. Tian, Systematic profiling of poly(A)+ transcripts modulated by core 3' end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation, PLoS Genet 11(4) (2015) e1005166.
[14] M. Chiabudini, C. Conz, F. Reckmann, S. Rospert, Ribosome-associated complex and Ssb are required for translational repression induced by polylysine segments within nascent chains, Mol Cell Biol 32(23) (2012) 4769-79.
[15] S. Danckwardt, M.W. Hentze, A.E. Kulozik, 3' end mRNA processing: molecular mechanisms and implications for health and disease, EMBO J 27(3) (2008) 482-98.
[16] R.M. Denome, C.N. Cole, Patterns of polyadenylation site selection in gene constructs containing multiple polyadenylation signals, Mol Cell Biol 8(11) (1988) 4829-39.
[17] B. Gawande, M.D. Robida, A. Rahn, R. Singh, Drosophila Sex-lethal protein mediates polyadenylation switching in the female germline, EMBO J 25(6) (2006) 1263-72.
[18] Y. Shi, Alternative polyadenylation: new insights from global analyses, RNA 18(12) (2012) 2105-17.
[19] P.A. Pinto, T. Henriques, M.O. Freitas, T. Martins, R.G. Domingues, P.S. Wyrzykowska, P.A. Coelho, A.M. Carmo, C.E. Sunkel, N.J. Proudfoot, A. Moreira, RNA polymerase II kinetics in polo polyadenylation signal selection, EMBO J 30(12) (2011) 2431-44.
[20] Z. Ji, W. Luo, W. Li, M. Hoque, Z. Pan, Y. Zhao, B. Tian, Transcriptional activity regulates alternative cleavage and polyadenylation, Mol Syst Biol 7 (2011) 534.
[21] A. Swaminathan, T.H. Beilharz, Epitope-tagged yeast strains reveal promoter driven changes to 3'-end formation and convergent antisense-transcription from common 3' UTRs, Nucleic Acids Res 44(1) (2016) 377-86.

[22] M.L. Peterson, S. Bertolino, F. Davis, An RNA Polymerase Pause Site Is Associated with the Immunoglobulin s Poly(A) Site, Molecular and Cellular Biology 22(15) (2002) 5606-5615.
[23] T.N. Mavrich, I.P. Ioshikhes, B.J. Venters, C. Jiang, L.P. Tomsho, J. Qi, S.C. Schuster, I. Albert, B.F. Pugh, A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome, Genome Res 18(7) (2008) 1073-83.
[24] S. Shivaswamy, A. Bhinge, Y. Zhao, S. Jones, M. Hirst, V.R. Iyer, Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation, PLoS Biol 6(3) (2008) e65.
[25] N. Spies, C.B. Nielsen, R.A. Padgett, C.B. Burge, Biased chromatin signatures around polyadenylation sites and exons, Mol Cell 36(2) (2009) 245-54.
[26] W. Li, J.Y. Park, D. Zheng, M. Hoque, G. Yehia, B. Tian, Alternative cleavage and polyadenylation in spermatogenesis connects chromatin regulation with post-transcriptional control, BMC Biol 14 (2016) 6.
[27] A.J. Wood, R. Schulz, K. Woodfine, K. Koltowska, C.V. Beechey, J. Peters, D. Bourc'his, R.J. Oakey, Regulation of alternative polyadenylation by genomic imprinting, Genes Dev 22(9) (2008) 1141-6.
[28] N.J. Proudfoot, G.G. Brownlee, 3' Non-coding region seqeunces in eukaryotic messenger RNA, Nature 263 (1976) 211-4.
[29] G. Martin, A.R. Gruber, W. Keller, M. Zavolan, 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(6) (2012) 753-63.
[30] M.I. Zarudnaya, I.M. Kolomiets, A.L. Potyahaylo, D.M. Hovorun, Downstream elements of mammalian pre-mRNA polyadenylation signals: primary, secondary and higher-order structures, Nucleic Acids Res 31(5) (2003) 1375-86.
[31] M. Mangone, A.P. Manoharan, D. Thierry-Mieg, J. Thierry-Mieg, T. Han, S.D. Mackowiak, E. Mis, C. Zegar, M.R. Gutwein, V. Khivansara, O. Attie, K. Chen, K. Salehi-Ashtiani, M. Vidal, T.T. Harkins, P. Bouffard, Y. Suzuki, S. Sugano, Y. Kohara, N. Rajewsky, F. Piano, K.C. Gunsalus, J.K. Kim, The landscape of C. elegans 3'UTRs, Science 329(5990) (2010) 432-5.
[32] N.J. Proudfoot, Ending the message: poly(A) signals then and now, Genes Dev 25(17) (2011) 1770-82.
[33] N.M. Nunes, W. Li, B. Tian, A. Furger, A functional human Poly(A) site requires only a potent DSE and an A-rich upstream sequence, EMBO J 29(9) (2010) 1523-36.
[34] B. Tian, J. Hu, H. Zhang, C.S. Lutz, A large-scale analysis of mRNA polyadenylation of human and mouse genes, Nucleic Acids Res 33(1) (2005) 201-12.
[35] E. Beaudoing, S. Freier, J.R. Wyatt, J.M. Claverie, D. Gautheret, Patterns of variant polyadenylation signal usage in human genes, Genome Res 10(7) (2000) 1001-10.
[36] A. Gil, N.J. Proudfoot, Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit ß-globin mRNA 3'end formation, Cell 49 (1987) 399-406.
[37] J. Hu, C.S. Lutz, J. Wilusz, B. Tian, Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation, RNA 11(10) (2005) 1485-93.
[38] J.H. Graber, C.R. Cantor, S.C. Mohr, T.F. Smith, In silico detection of control signals: mRNA 3'-end-processing sequences in diverse species, Proc Natl Acad Sci U S A 96(24) (1999) 14055-60.
[39] K. Venkataraman, K.M. Brown, G.M. Gilmartin, Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition, Genes Dev 19(11) (2005) 1315-27.
[40] C. Yao, J. Biesinger, J. Wan, L. Weng, Y. Xing, X. Xie, Y. Shi, Transcriptome-wide analyses of CstF64-RNA interactions in global regulation of mRNA alternative polyadenylation, Proc Natl Acad Sci U S A 109(46) (2012) 18773-8.

[41] J. Zhao, L. Hyman, C. Moore, Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis, Microbiol Mol Biol Rev 63(2) (1999) 405-45.
[42] Y. Shi, D.C. Di Giammartino, D. Taylor, A. Sarkeshik, W.J. Rice, J.R. Yates, 3rd, J. Frank, J.L. Manley, Molecular architecture of the human pre-mRNA 3' processing complex, Mol Cell 33(3) (2009) 365-76.
[43] C.R. Mandel, Y. Bai, L. Tong, Protein factors in pre-mRNA 3'-end processing, Cell Mol Life Sci 65(7-8) (2008) 1099-122.
[44] Y. Takagaki, L.C. Ryner, J.L. Manley, Four factors are require for 3'-end cleavage of pre-mRNAs, Genes Dev 3 (1989) 1711-24.
[45] S. Chan, E.A. Choi, Y. Shi, Pre-mRNA 3'-end processing complex assembly and function, Wiley Interdiscip Rev RNA 2(3) (2011) 321-35.
[46] D.F. Colgan, J.L. Manley, Mechanism and regulation of mRNA polyadenylation, Genes Dev 11(21) (1997) 2755-66.
[47] S. Millevoi, S. Vagner, Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation, Nucleic Acids Res 38(9) (2010) 2757-74.
[48] K. Xiang, L. Tong, J.L. Manley, Delineating the structural blueprint of the pre-mRNA 3'-end processing machinery, Mol Cell Biol 34(11) (2014) 1894-910.
[49] Y. Takagaki, R.L. Seipelt, M.L. Peterson, J.L. Manley, The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation, Cell 87(5) (1996) 941-52.
[50] Y. Takagaki, J.L. Manley, Levels of polyadenylation factor CstF-64 control IgM heavy chain mRNA accumulation and other events associated with B cell differentiation, Mol Cell 2(6) (1998) 761-71.
[51] Z. Xia, L.A. Donehower, T.A. Cooper, J.R. Neilson, D.A. Wheeler, E.J. Wagner, W. Li, Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types, Nat Commun 5 (2014) 5274.
[52] C. Yao, E.A. Choi, L. Weng, X. Xie, J. Wan, Y. Xing, J.J. Moresco, P.G. Tu, J.R. Yates, 3rd, Y. Shi, Overlapping and distinct functions of CstF64 and CstF64tau in mammalian mRNA 3' processing, RNA 19(12) (2013) 1781-90.
[53] B. Lackford, C. Yao, G.M. Charles, L. Weng, X. Zheng, E.A. Choi, X. Xie, J. Wan, Y. Xing, J.M. Freudenberg, P. Yang, R. Jothi, G. Hu, Y. Shi, Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal, EMBO J 33(8) (2014) 878-89.
[54] K.S.K. Guisbert, H. Li, C. Guthrie, Alternative 3' pre-mRNA processing in Saccharomyces cerevisiae is modulated by Nab4/Hrp1 in vivo, PLoS Biol 5(1) (2007) e6.
[55] T. Kubo, T. Wada, Y. Yamaguchi, A. Shimizu, H. Handa, Knock-down of 25 kDa subunit of cleavage factor Im in Hela cells alters alternative polyadenylation within 3'-UTRs, Nucleic Acids Res 34(21) (2006) 6264-71.
[56] S. Kim, J. Yamamoto, Y. Chen, M. Aida, T. Wada, H. Handa, Y. Yamaguchi, Evidence that cleavage factor Im is a heterotetrameric protein complex controlling alternative polyadenylation, Genes Cells 15(9) (2010) 1003-13.
[57] A.R. Gruber, G. Martin, W. Keller, M. Zavolan, Cleavage factor Im is a key regulator of 3' UTR length, RNA Biol 9(12) (2012) 1405-12.
[58] C.P. Masamha, Z. Xia, J. Yang, T.R. Albrecht, M. Li, A.B. Shyu, W. Li, E.J. Wagner, CFIm25 links alternative polyadenylation to glioblastoma tumour suppression, Nature 510(7505) (2014) 412-6.
[59] M. Jenal, R. Elkon, F. Loayza-Puch, G. van Haaften, U. Kuhn, F.M. Menzies, J.A. Oude Vrielink, A.J. Bos, J. Drost, K. Rooijers, D.C. Rubinsztein, R. Agami, The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites, Cell 149(3) (2012) 538-53.

[60] S. McCracken, N. Fong, K. Yankulov, S. Ballantyne, G. Pan, J. Greenblatt, S.D. Patterson, M. Wickens, D.L. Bentley, The C-terminal domain of RNA polymerase II couples mRNA processing to transcription, Nature 385(6614) (1997) 357-61.
[61] Y. Hirose, J.L. Manley, RNA polymerase II is an essential mRNA polyadenylation factor, Nature 395 (1998) 93-6.
[62] K. Glover-Cutter, RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes, Nat Struct Mol Biol 15(1) (2008) 71-8.
[63] D.D. Licatalosi, G. Geiger, M. Minet, S. Schroeder, K. Cilli, J.B. McNeil, D.L. Bentley, Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II, Mol Cell 9(5) (2002) 1101-11.
[64] B. Dichtl, D. Blank, M. Sadowski, W. Hubner, S. Weiser, W. Keller, Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination, EMBO J 21(15) (2002) 4125-35.
[65] M. Sadowski, B. Dichtl, W. Hubner, W. Keller, Independent functions of yeast Pcf11p in pre-mRNA 3'end processing and in transcription termination, EMBO J 22(9) (2003) 2167-77.
[66] T. Nagaike, C. Logan, I. Hotta, O. Rozenblatt-Rosen, M. Meyerson, J.L. Manley, Transcriptional activators enhance polyadenylation of mRNA precursors, Mol Cell 41(4) (2011) 409-18.
[67] Y. Yang, W. Li, M. Hoque, L. Hou, S. Shen, B. Tian, B.D. Dynlacht, PAF Complex Plays Novel Subunit-Specific Roles in Alternative Cleavage and Polyadenylation, PLoS Genet 12(1) (2016) e1005794.
[68] K. Martincic, S.A. Alkan, A. Cheatle, L. Borghesi, C. Milcarek, Transcription elongation factor ELL2 directs immunoglobulin secretion in plasma cells by stimulating altered RNA processing, Nat Immunol 10(10) (2009) 1102-9.
[69] D. Kaida, M.G. Berg, I. Younis, M. Kasim, L.N. Singh, L. Wan, G. Dreyfuss, U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation, Nature 468(7324) (2010) 664-8.
[70] M.G. Berg, L.N. Singh, I. Younis, Q. Liu, A.M. Pinto, D. Kaida, Z. Zhang, S. Cho, S. Sherrill-Mix, L. Wan, G. Dreyfuss, U1 snRNP determines mRNA length and regulates isoform expression, Cell 150(1) (2012) 53-64.
[71] A. Kyburz, A. Friedlein, L. H., W. Keller, 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 (2006) 195-205.
[72] S. Millevoi, C. Loulergue, S. Dettwiler, S.Z. Karaa, W. Keller, M. Antoniou, S. Vagner, An interaction between U2AF 65 and CFIm links the splicing and 3' end processing machineries, EMBO J 25 (2006) 4854-64.
[73] M. Nazim, A. Masuda, M.A. Rahman, F. Nasrin, J.I. Takeda, K. Ohe, B. Ohkawara, M. Ito, K. Ohno, Competitive regulation of alternative splicing and alternative polyadenylation by hnRNP H and CstF64 determines acetylcholinesterase isoforms, Nucleic Acids Res 45(3) (2017) 1455-1468.
[74] M. Movassat, T.L. Crabb, A. Busch, C. Yao, D.J. Reynolds, Y. Shi, K.J. Hertel, Coupling between alternative polyadenylation and alternative splicing is limited to terminal introns, RNA Biol 13(7) (2016) 646-55.
[75] P. Castelo-Branco, A. Furger, M. Wollerton, C. Smith, A. Moreira, N. Proudfoot, Polypyrimidine tract binding protein modulates efficiency of polyadenylation, Mol Cell Biol 24(10) (2004) 4174-83.
[76] R. Batra, K. Charizanis, M. Manchanda, A. Mohan, M. Li, D.J. Finn, M. Goodwin, C. Zhang, K. Sobczak, C.A. Thornton, M.S. Swanson, Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease, Mol Cell 56(2) (2014) 311-22.

[77] U. Kuhn, M. Gundel, A. Knoth, Y. Kerwitz, S. Rudel, E. Wahle, Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor, J Biol Chem 284(34) (2009) 22803-14.
[78] J. Katahira, D. Okuzaki, H. Inoue, Y. Yoneda, K. Maehara, Y. Ohkawa, Human TREX component Thoc5 affects alternative polyadenylation site choice by recruiting mammalian cleavage factor I, Nucleic Acids Res 41(14) (2013) 7060-72.
[79] F.A. Bava, C. Eliscovich, P.G. Ferreira, B. Minana, C. Ben-Dov, R. Guigo, J. Valcarcel, R. Mendez, CPEB1 coordinates alternative 3'-UTR formation with translational regulation, Nature 495(7439) (2013) 121-5.
[80] S. Danckwardt, A.S. Gantzert, S. Macher-Goeppinger, H.C. Probst, M. Gentzel, M. Wilm, H.J. Grone, P. Schirmacher, M.W. Hentze, A.E. Kulozik, p38 MAPK controls prothrombin expression by regulated RNA 3' end processing, Mol Cell 41(3) (2011) 298-310.
[81] V. Hilgers, S.B. Lemke, M. Levine, ELAV mediates 3' UTR extension in the Drosophila nervous system, Genes Dev 26(20) (2012) 2259-64.
[82] K.D. Mansfield, J.D. Keene, Neuron-specific ELAV/Hu proteins suppress HuR mRNA during neuronal differentiation by alternative polyadenylation, Nucleic Acids Res 40(6) (2012) 2734-46.
[83] W. Dai, G. Zhang, E.V. Makeyev, RNA-binding protein HuR autoregulates its expression by promoting alternative polyadenylation site usage, Nucleic Acids Res 40(2) (2012) 787-800.
[84] M. Samson, Evidence for 3'untranslated region-dependent autoregulation of the Drosophila gene encoding the neuronal nuclear RNA-binding protein ELAV, Genetics 150 (1998) 723-33.
[85] W. Luo, Z. Ji, Z. Pan, B. You, M. Hoque, W. Li, S.I. Gunderson, B. Tian, The conserved intronic cleavage and polyadenylation site of CstF-77 gene imparts control of 3' end processing activity through feedback autoregulation and by U1 snRNP, PLoS Genet 9(7) (2013) e1003613.
[86] B.R. Graverley, E.S. Fleming, G.M. Gilmartin, RNA structure is a critical determinant of poly(A) site recognition by cleavage and polyadenylation specificity factor, Mol Cell Biol 16(9) (1996) 4942-51.
[87] M.J. Munoz, R.R. Daga, A. Garzon, G. Thode, J. Jimenez, Poly(A) site choice during mRNA 3'-end formation in the Schizosaccharomyces pombe wos2 gene, Mol Genet Genomics 267(6) (2002) 792-6.
[88] R. Flomen, A. Makoff, Increased RNA editing in EAAT2 pre-mRNA from amyotrophic lateral sclerosis patients: involvement of a cryptic polyadenylation site, Neurosci Lett 497(2) (2011) 139-43.
[89] R.S. Redis, L.E. Vela, W. Lu, J. Ferreira de Oliveira, C. Ivan, C. Rodriguez-Aguayo, D. Adamoski, B. Pasculli, A. Taguchi, Y. Chen, A.F. Fernandez, L. Valledor, K. Van Roosbroeck, S. Chang, M. Shah, G. Kinnebrew, L. Han, Y. Atlasi, L.H. Cheung, G.Y. Huang, P. Monroig, M.S. Ramirez, T. Catela Ivkovic, L. Van, H. Ling, R. Gafa, S. Kapitanovic, G. Lanza, J.A. Bankson, P. Huang, S.Y. Lai, R.C. Bast, M.G. Rosenblum, M. Radovich, M. Ivan, G. Bartholomeusz, H. Liang, M.F. Fraga, W.R. Widger, S. Hanash, I. Berindan-Neagoe, G. Lopez-Berestein, A.L. Ambrosio, S.M. Gomes Dias, G.A. Calin, Allele-Specific Reprogramming of Cancer Metabolism by the Long Non-coding RNA CCAT2, Mol Cell 61(4) (2016) 520-34.
[90] A.R. Gruber, G. Martin, P. Muller, A. Schmidt, A.J. Gruber, R. Gumienny, N. Mittal, R. Jayachandran, J. Pieters, W. Keller, E. van Nimwegen, M. Zavolan, Global 3' UTR shortening has a limited effect on protein abundance in proliferating T cells, Nat Commun 5 (2014) 5465.
[91] E.E. Creemers, A. Bawazeer, A.P. Ugalde, H.W. van Deutekom, I. van der Made, N.E. de Groot, M.E. Adriaens, S.A. Cook, C.R. Bezzina, N. Hubner, J. van der Velden, R. Elkon, R. Agami,

Y.M. Pinto, Genome-Wide Polyadenylation Maps Reveal Dynamic mRNA 3'-End Formation in the Failing Human Heart, Circ Res 118(3) (2016) 433-8.
[92] Y. Fu, Y. Sun, Y. Li, J. Li, X. Rao, C. Chen, A. Xu, Differential genome-wide profiling of tandem 3' UTRs among human breast cancer and normal cells by high-throughput sequencing, Genome Res 21(5) (2011) 741-7.
[93] R. Sandberg, J.R. Neilson, A. Sarma, P.A. Sharp, C.B. Burge, Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites, Science 320(5883) (2008) 1643-7.
[94] C. Mayr, D.P. Bartel, Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells, Cell 138(4) (2009) 673-84.
[95] L. Li, D. Wang, M. Xue, X. Mi, Y. Liang, P. Wang, 3'UTR shortening identifies high-risk cancers with targeted dysregulation of the ceRNA network, Sci Rep 4 (2014) 5406.
[96] A.R. Gruber, G. Martin, W. Keller, M. Zavolan, Means to an end: mechanisms of alternative polyadenylation of messenger RNA precursors, Wiley Interdiscip Rev RNA 5(2) (2014) 183-96.
[97] C. Desmedt, B. Haibe-Kains, P. Wirapati, M. Buyse, D. Larsimont, G. Bontempi, M. Delorenzi, M. Piccart, C. Sotiriou, Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes, Clin Cancer Res 14(16) (2008) 5158-65.
[98] J.S. Parker, M. Mullins, M.C. Cheang, S. Leung, D. Voduc, T. Vickery, S. Davies, C. Fauron, X. He, Z. Hu, J.F. Quackenbush, I.J. Stijleman, J. Palazzo, J.S. Marron, A.B. Nobel, E. Mardis, T.O. Nielsen, M.J. Ellis, C.M. Perou, P.S. Bernard, Supervised risk predictor of breast cancer based on intrinsic subtypes, J Clin Oncol 27(8) (2009) 1160-7.
[99] S. Lianoglou, V. Garg, J.L. Yang, C.S. Leslie, C. Mayr, Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression, Genes Dev 27(21) (2013) 2380-96.
[100] C. Mayr, Evolution and Biological Roles of Alternative 3'UTRs, Trends Cell Biol 26(3) (2016) 227-37.
[101] B.D. Berkovits, C. Mayr, Alternative 3' UTRs act as scaffolds to regulate membrane protein localization, Nature 522(7556) (2015) 363-7.
[102] N. Spies, C.B. Burge, D.P. Bartel, 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts, Genome Res 23(12) (2013) 2078-90.
[103] Y. Yoon, M.C. McKenna, D.A. Rollins, M. Song, T. Nuriel, S.S. Gross, G. Xu, C.E. Glatt, Anxiety-associated alternative polyadenylation of the serotonin transporter mRNA confers translational regulation by hnRNPK, Proc Natl Acad Sci U S A 110(28) (2013) 11624-9.
[104] P. Oikonomou, H. Goodarzi, S. Tavazoie, Systematic identification of regulatory elements in conserved 3' UTRs of human transcripts, Cell Rep 7(1) (2014) 281-92.
[105] J.R. Dickson, C. Kruse, D.R. Montagna, B. Finsen, M.S. Wolfe, Alternative polyadenylation and miR-34 family members regulate tau expression, J Neurochem 127(6) (2013) 739-49.
[106] H. Rhinn, L. Qiang, T. Yamashita, D. Rhee, A. Zolin, W. Vanti, A. Abeliovich, Alternative alpha-synuclein transcript usage as a convergent mechanism in Parkinson's disease pathology, Nat Commun 3 (2012) 1084.
[107] V.A. Gennarino, C.E. Alcott, C.A. Chen, A. Chaudhury, M.A. Gillentine, J.A. Rosenfeld, S. Parikh, J.W. Wheless, E.R. Roeder, D.D. Horovitz, E.K. Roney, J.L. Smith, S.W. Cheung, W. Li, J.R. Neilson, C.P. Schaaf, H.Y. Zoghbi, NUDT21-spanning CNVs lead to neuropsychiatric disease and altered MeCP2 abundance via alternative polyadenylation, Elife 4 (2015).
[108] R. Soetanto, C.J. Hynes, H.R. Patel, D.T. Humphreys, M. Evers, G. Duan, B.J. Parker, S.K. Archer, J.L. Clancy, R.M. Graham, T.H. Beilharz, N.J. Smith, T. Preiss, Role of miRNAs and alternative mRNA 3'-end cleavage and polyadenylation of their mRNA targets in cardiomyocyte hypertrophy, Biochim Biophys Acta 1859(5) (2016) 744-56.

[109] M.D. Barnhart, S.L. Moon, A.W. Emch, C.J. Wilusz, J. Wilusz, Changes in cellular mRNA stability, splicing, and polyadenylation through HuR protein sequestration by a cytoplasmic RNA virus, Cell Rep 5(4) (2013) 909-17.
[110] S. Tang, A. Patel, P.R. Krause, Herpes simplex virus ICP27 regulates alternative pre-mRNA polyadenylation and splicing in a sequence-dependent manner, Proc Natl Acad Sci U S A 113(43) (2016) 12256-12261.
[111] W.O. Miles, A. Lembo, A. Volorio, E. Brachtel, B. Tian, D. Sgroi, P. Provero, N. Dyson, Alternative Polyadenylation in Triple-Negative Breast Tumors Allows NRAS and c-JUN to Bypass PUMILIO Posttranscriptional Regulation, Cancer Res 76(24) (2016) 7231-7241.

FIGURE LEGENDS

Figure 1. Mechanisms of Alternative Polyadenylation

Polyadenylation events can be split into five different categories, four of which involve alternative polyadenylation. **A)** Constitutive polyadenylation involves the occurrence of a single potential poly(A) site within the 3' UTR of the transcript. **B)** Tandem 3'UTR APA genes possess two or more cleavage poly(A) sites within their 3'UTR. Resulting transcripts only differ by the length of their untranslated region. **C)** Alternative terminal exon APA requires alternative splicing to occur that changes the last exon and therefore the available poly(A) site. **D)** Intronic APA involves the use of cryptic alternative poly(A) sites found within introns. **E)** Internal exon APA uses poly(A) sites within upstream exons and results in a transcript lacking a stop codon or a 3'UTR.

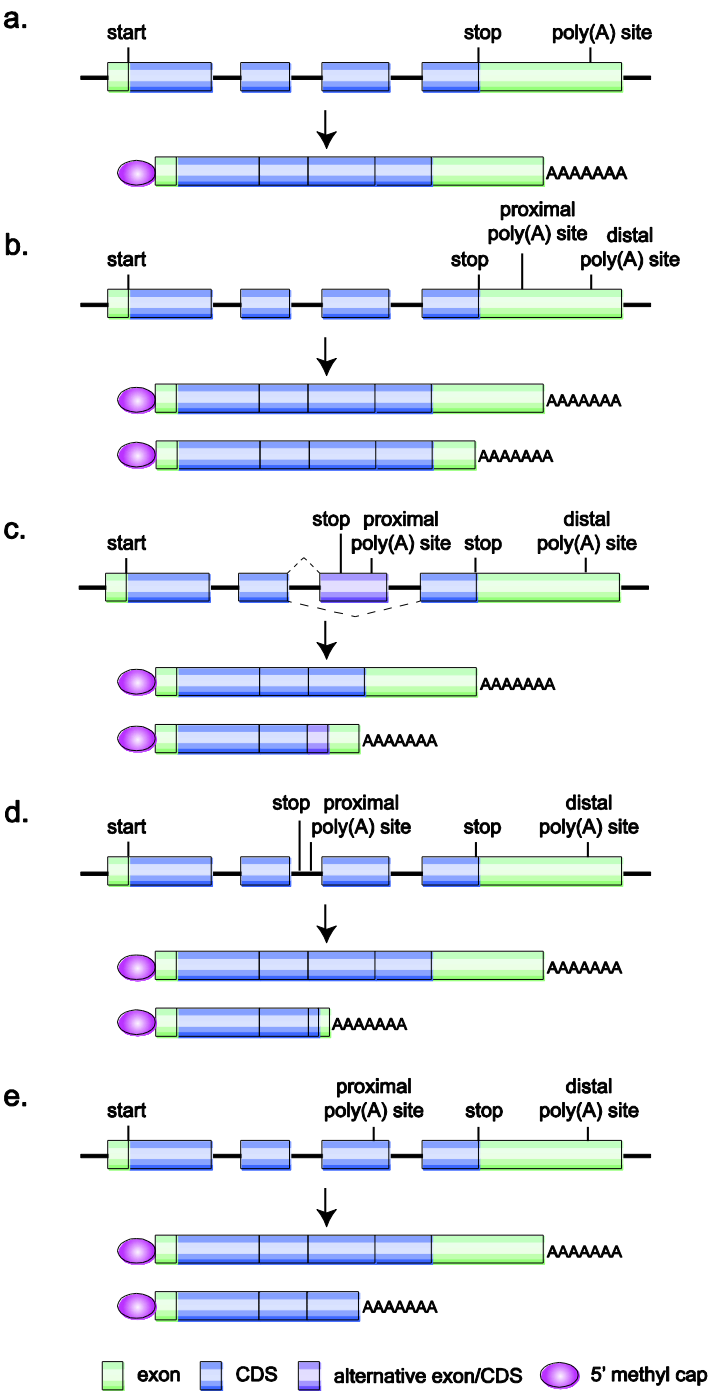


Figure 2. Core Cis Elements Involved in Cleavage and Polyadenylation Site Recognition

The cleavage and polyadenylation machinery is guided to target RNA by a series of *cis* elements as indicated. These are generalised here in schematic form, but it is important to note that few given genes will encode all elements. The positioning elements in simple eukaryotes such as *S. cerevisiae* tend to be more divergent than their mammalian counterparts.

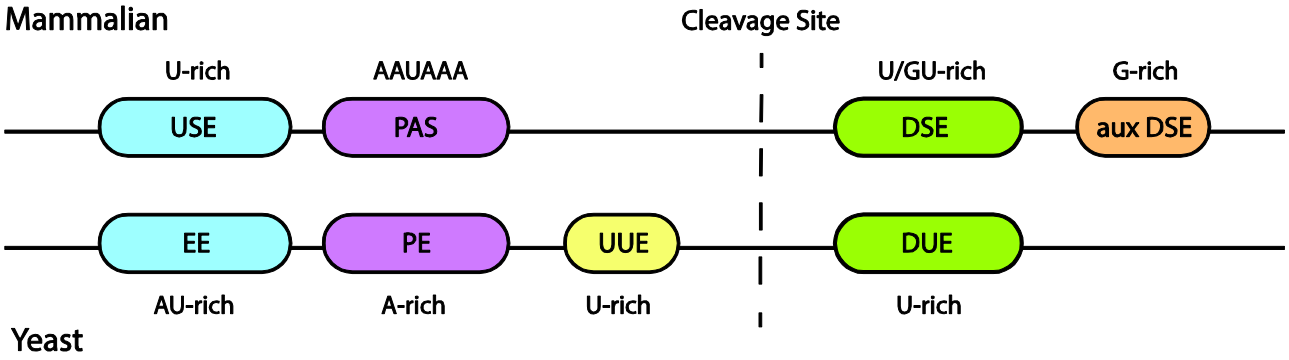
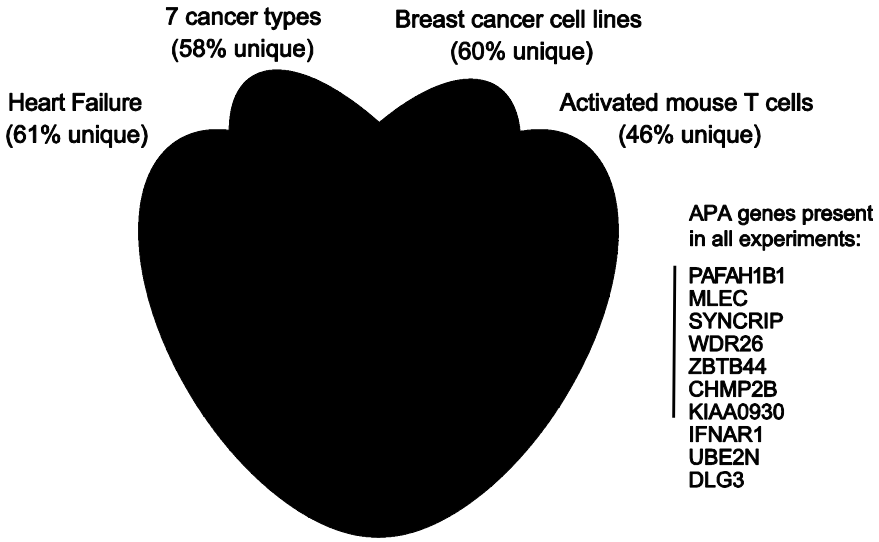


Figure 3. APA is disease and condition specific

Genome-wide APA genes are more likely to be unique than shared among disease states. **A)** Venn-diagram of overlapping gene sets from 4 different experimental datasets. The APA events determined by comparing test versus control samples from the failing human heart [91], 7 cancer types [51],

2 breast cancer cell lines (HER2 positive and negative) [92] and activated mouse T cells [90]. Only 10 genes were shared in APA genes in all experiments and these are listed to the right of the diagram. **B)** Venn-diagram of overlapping APA events from 5 of the 7 tumour types inferred from RNA-seq coverage [51]. The percentage of APA genes unique to each condition are indicated.

a.



b.

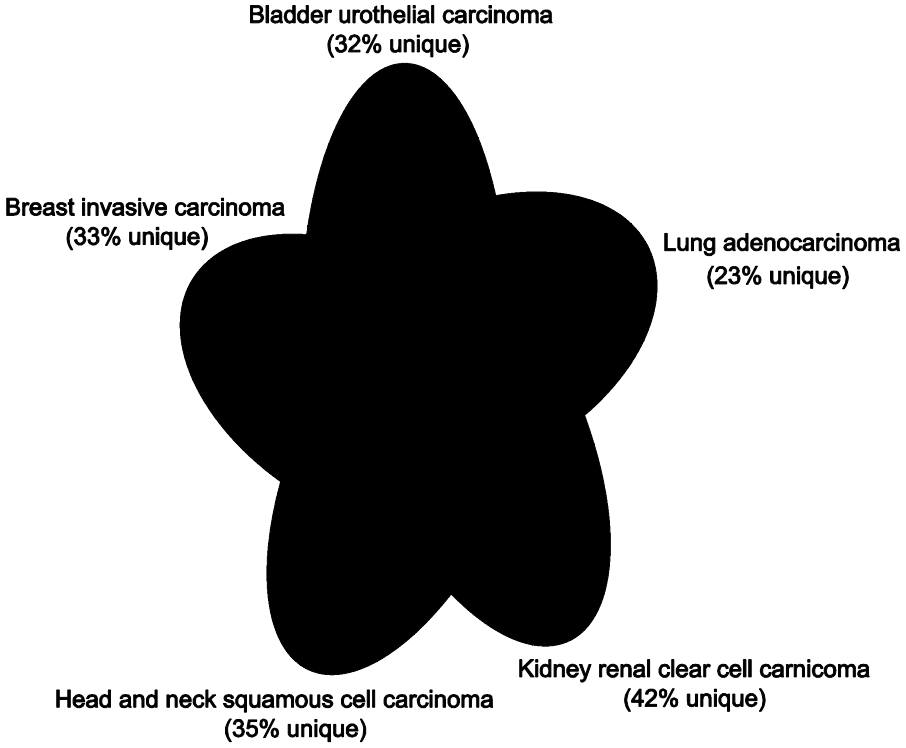


Table 1. The core cleavage and polyadenylation factors in yeast and mammalian systems.

Mammalian Factor	**Mammalian Complex**	**Yeast Homologue Factor**	**Yeast Complex**
Symplekin		Pta1	CPF(CFII)
CPSF-160 (CPSF1)	CPSF	Cft1 (Yhh1)	CPF(CFII)
CPSF-100 (CPSF2)	CPSF	Cft2 (Ydh1)	CPF(CFII)
CPSF-73 (CPSF3)	CPSF	Ysh1 (Brr5)	CPF(CFII)
CPSF-30 (CPSF4)	CPSF	Yth1	CPF(PFI)
Fip1 (FIP1L1)	CPSF	Fip1	CPF(PFI)
Wdr33	CPSF	Pfs2	CPF(PFI)
PP1		Glc7	CPF
Pap (PAPOLA)		Pap1	CPF
Rbbp6		Mpe1	CPF
		Pti1	
		Ref2	CPF
Ssu72		Ssu72	CPF
Wdr82		Swd2	
		Syc1	
CstF-50 (CSTF1)	CstF		
CstF-64 (CSTF2)	CstF	Rna15	CFIA
CstF-77 (CSTF3)	CstF	Rna14	CFIA
CFIm-25 (CPSF5/NUDT21)	CFIm		
CFIm-68 (CPSF6)	CFIm		
CFIm-59 (CPSF7)	CFIm		
Pcf11	CFIIm	Pcf11	CFIA
Clp1	CFIIm	Clp1	CFIA
		Hrp1 (Nab4)	CFIB

Table 2. Recent studies with disease associated APA events.

Disease	**Summary**
Alzheimer's disease (AD)	Switch to short Tau 3' UTR isoform in AD patients. Loss of miR-34a repression leads to more aggregates [105]
Parkinson's disease (PD)	Switch to long α-Synuclein 3'UTR isoforms increases protein expression and formation of Lewy bodies in PD [106]
Myotonic dystrophy (DM)	Misregulated APA in DM by inhibition of MBNL proteins [76]
Neuropsychiatric disease	Increased CFIm-25 causes switch to distal cleavage site in the MECP2 3'UTR and neuropsychiatric disease [107]
Heart failure	Altered APA of a subset of genes in the failing human heart [91]
Cardiac hypertrophy	Global trend toward 3' UTR shortening in cardiac hypertrophy [3, 108]
Sindbis virus	Infection causes redistribution of HuR protein to the cytoplasm and altered HuR-regulated APA events [109]
Herpes simplex virus (HSV)	Infected cell culture polypeptide 27 (ICP27) promotes cryptic APA in host cells [110]
Glioblastoma	CFIm-25 depletion causes switch to proximal 3' UTRs in glioblastoma tumours [58]
Pan cancer APA	Global 3' UTR shortening associated with tumorigenesis and CstF-64 proposed as master regulator of APA [51]
Prostate cancer	APA in prostate cancer changed the availability of miRNA binding sites, modulating competing endogenous RNA (ceRNA) networks [95]
Triple negative breast cancer (TNBC)	RNA binding protein sites are lost in TNBC through 3' UTR shortening [111]