Published OnlineFirst April 13, 2016; DOI: 10.1158/1541-7786.MCR-15-0489

Review

Molecular Cancer Research

Alternative Polyadenylation: Another Foe in Cancer

Ayse Elif Erson-Bensan[1] and Tolga Can[2]

Abstract

Advancements in sequencing and transcriptome analysis methods have led to seminal discoveries that have begun to unravel the complexity of cancer. These studies are paving the way toward the development of improved diagnostics, prognostic predictions, and targeted treatment options. However, it is clear that pieces of the cancer puzzle are still missing. In an effort to have a more comprehensive understanding of the development and progression of cancer, we have come to appreciate the value of the noncoding regions of our genomes, partly due to the discovery of miRNAs and their significance in gene regulation. Interestingly, the miRNA–mRNA interactions are not solely dependent on variations in miRNA levels. Instead, the majority of genes harbor multiple polyadenylation signals on their 3′ UTRs (untranslated regions) that can be differentially selected on the basis of the physiologic state of cells, resulting in alternative 3′ UTR isoforms. Deregulation of alternative polyadenylation (APA) has increasing interest in cancer research, because APA generates mRNA 3′ UTR isoforms with potentially different stabilities, subcellular localizations, translation efficiencies, and functions. This review focuses on the link between APA and cancer and discusses the mechanisms as well as the tools available for investigating APA events in cancer. Overall, detection of deregulated APA-generated isoforms in cancer may implicate some proto-oncogene activation cases of unknown causes and may help the discovery of novel cases; thus, contributing to a better understanding of molecular mechanisms of cancer. *Mol Cancer Res; 14(6); 507–17. ©2016 AACR.*

End of the Message Matters

The 3′ untranslated regions (UTR) have long been known to have important roles in maintaining the stability, localization, and half-life of the mRNAs but a detailed mechanistic explanation as to how these properties are regulated or the consequences of deregulation are just beginning to be appreciated. Except for the replication-dependent histones processed by small nuclear ribonucleoproteins (1), all eukaryotic mRNA 3′ UTRs undergo endonucleolytic cleavage by polyadenylation machinery proteins and polyadenylation by the poly(A) polymerases (PAP; refs. 2, 3). Therefore, the position of the poly(A) signal is one of the main determinants of where the pre-mRNA will be cleaved and "tailed" at the poly(A) site. Poly(A) signals are usually found at approximately 15–30 nucleotides upstream regions of the actual cleavage sites. The canonical poly(A) signal sequence is AAUAAA and is strongly enriched and conserved among mammalian species, including humans. There are also poly(A) signal variants (e.g., AAAUAA, AUAAAA, AUUAAA, AUAAAU, AUAAAG, CAAUAA, UAAUAA, AUAAAC, AAAAUA, AAAAAA, AAAAAG) that are utilized with varying frequencies throughout the genome (4).

In addition to the poly(A) signal, uridine (U)-rich elements (*i.e.*, UGUA) at approximately 10–35 nucleotides upstream and U-/GU-rich regions downstream the poly(A) signal help to position the polyadenylation machinery (5, 6). The exact functional significance of these conserved sites or existence of other variants require further investigation as 20% of human poly(A) signals are not surrounded by U-/GU-rich regions (7). However, it is known that the recognition of the poly(A) signal, cleavage, and polyadenylation is orchestrated by a multi-protein machinery consisting of three main complexes; (i) cleavage and polyadenylation stimulatory factor (CPSF), (ii) cleavage stimulatory factor (CSTF), and (iii) cleavage factors Im and IIm (CFIm, CFIIm). Symplekin functions as a scaffold protein within this multi-protein complex. Poly(A)-binding proteins (PABP) stimulate PAPs to catalyze the addition of untemplated adenosines and regulate the length of the poly(A) tails (8). Known key polyadenylation proteins are summarized below and in Fig. 1.

CPSF recognizes the polyadenylation signal AAUAAA, providing sequence specificity. The complex consists of hFip1, WDR33 (WD repeat domain 33), CPSF30, CPSF73, CPSF100, and CPSF160 subunits (9, 10). Although conflicting results were reported for whether CPSF30 or CPSF160 binds to AAUAAA, WDR33 is indeed the subunit that binds the AAUAAA signal (11, 12). CPSF73 is the endonuclease that cleaves the pre-mRNA at the poly(A) site (13). The CPSF complex also interacts with TFIID (transcription factor II D), RNA polymerase II (14–16), and spliceosomal factors (17). On the other hand, the CSTF complex binds as a dimer to the downstream region of the cleavage site and contains CSTF1 (50 kDa), CSTF2 (64 kDa), and CSTF3 (77 kDa; refs. 5, 18). The CSTF complex is essential for the cleavage reaction (19). While CSTF2 (also known as CSTF64) interacts directly with the distal sequence elements (GU-rich region; refs. 20, 21), CSTF1 and CSTF3 subunits interact with the C-terminal domain of RNA polymerase II, possibly to facilitate activation of other 3′ UTR processing proteins (22).

[1]Department of Biological Sciences, Middle East Technical University (METU) (ODTU), Ankara, Turkey. [2]Department of Computer Engineering, Middle East Technical University (METU) (ODTU), Ankara, Turkey.

Corresponding Author: Ayse Elif Erson-Bensan, Middle East Technical University, Dumlupinar Bulvar No: 1, Universiteler Mah., Ankara, Turkey Phone: 9031-2210-5043; Fax: 9031-2210-7976; E-mail: erson@metu.edu.tr

doi: 10.1158/1541-7786.MCR-15-0489

©2016 American Association for Cancer Research.

Downloaded from mcr.aacrjournals.org on July 24, 2020. © 2016 American Association for Cancer Research.

Published OnlineFirst April 13, 2016; DOI: 10.1158/1541-7786.MCR-15-0489

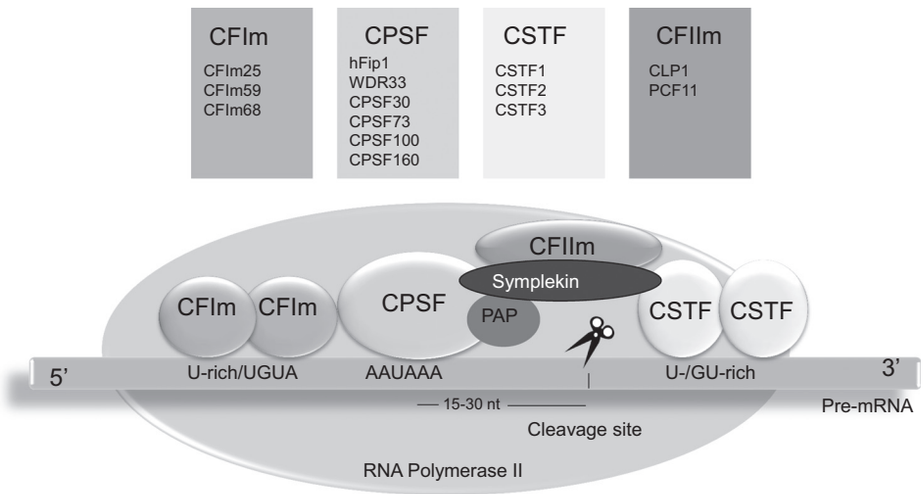


Figure 1.
Polyadenylation machinery. mRNA 3′ UTRs are recognized by three multi-protein complexes; cleavage and polyadenylation stimulatory factor (CPSF), cleavage stimulatory factor (CSTF), and cleavage factors Im and IIm (CFIm, CFIIm). CPSF complex consists of hFip1, WDR33 (WD repeat domain 33), CPSF30, CPSF73, CPSF100, and CPSF160. CPSF complex recognizes the polyadenylation signal AAUAAA and CPSF73 is the endonuclease that cleaves the mRNA. Symplekin functions as a scaffold protein and PAPs catalyze the addition of untemplated adenosines. CSTF complex binds to the downstream of the cleavage site and consists of CSTF1, CSTF2, and CSTF3. CFIm complex; CFIm25, CFIm59, and CFIm68, bind to upstream conserved elements to mediate recruitment of other proteins and the cleavage reaction. CFIIm complex; CLP1 (cleavage and polyadenylation factor I subunit 1) and PCF11 (cleavage and polyadenylation factor subunit) aid RNA polymerase II–mediated transcription termination.

The CFIm complex, consisting of CFIm25, CFIm59, and CFIm68, binds upstream conserved elements to mediate the cleavage reaction *in vitro* and *in vivo* (23–26). Multiple UGUA elements are generally present at 40–100 nucleotides upstream of the poly(A) site and CFIm25 binds to two UGUA elements as a homodimer (26). Furthermore, structural studies revealed that the two CFIm complexes bind the two UGUA elements in an antiparallel manner, resulting in looping of the mRNA (27, 28). Therefore, CFIm binding can function as a primary determinant of poly(A) site recognition by looping out an entire poly(A) site region and thereby inducing the selection of an alternative poly (A) site (28). Chromatin immunoprecipitation (ChIP) analyses indicate that CFIm is recruited to the transcription unit, along with CPSF and CSTF, as early as transcriptional initiation, supporting a direct role for CFIm in polyadenylation (24). The CFIIm complex, consisting of CLP1 (cleavage and polyadenylation factor I subunit 1) and PCF11 (cleavage and polyadenylation factor subunit), aids Pol II–mediated transcription termination (29, 30).

Overall, cleavage and polyadenylation are likely to be closely coupled where cleavage is immediately followed by polyadenylation suggesting that CPSF binding to the polyadenylation signal AAUAAA recruits PAPs (9, 31). PAPs are grouped as canonical and noncanonical polymerases. Cleaved mRNAs are generally polyadenylated by canonical PAPs, which are a small family of monomeric enzymes. Because of the low processivity of PAPs, both CPSF and the nuclear poly(A)-binding protein PABPN1 synergistically stimulate efficient polyadenylation and control the length of the poly(A) tail (reviewed in ref. 32). As a result, the poly(A) tail protects the mRNA against nucleases (reviewed in ref. 33). In addition, the length of the poly(A) tail is dynamically regulated with potential implications other than just protection from nucleases. For example, in early zebrafish and frog embryos, mean poly(A) tail lengths correlate strongly with translational efficiency. However, this correlation ends at the gastrulation stage indicating a switch in the mechanism of translational control at later developmental stages (34). An interesting finding showed that PAP activity itself is linked to poor prognosis in leukemia and breast cancer, perhaps hinting at the possibility that the poly(A) tail length is important in nonembryonic cells (31). Considering the existence of multiple alternatively spliced isoforms of PAPs (3), the functional significance of poly(A) tail length regulation remains to be further investigated in normal and cancer cells. On the other hand, noncanonical PAPs are cytoplasmic, where polyadenylation generally increases protein expression from specific mRNAs with short poly(A) tails. Initially identified in *C. elegans* and *S. pombe* and later in vertebrates and mammals, noncanonical PAPs and cytoplasmic polyadenylation have been implicated in many processes, including synaptic functions and mitosis (refs. 35–37; reviewed in ref. 38).

RBPs (RNA-binding proteins) can also affect cleavage and polyadenylation patterns of target mRNAs by competing with or enhancing the binding of the polyadenylation machinery proteins to their target sites. HNRNP H (heterogeneous nuclear ribonucleoprotein H) binds to proximal poly(A) sites and recruits the core polyadenylation machinery proteins resulting in 3′ UTR shortening events as demonstrated by high-throughput sequencing–crosslinking immunoprecipitation (HITS-CLIP). Similarly, the cytoplasmic polyadenylation element binding protein 1 (CPEB1) binds upstream of the weaker proximal poly(A) signal and recruits the CPSF complex to promote 3′ UTR shortening (39). MBNL (muscleblind-like) is another RBP that may positively or negatively affect the binding of polyadenylation proteins (e.g., CFIm68) to their target regions. Accordingly, depletion of Mbnl in mouse embryo fibroblasts leads to deregulation of thousands of alternative polyadenylation events (40). Alternatively, *Drosophila* ELAV (embryonic lethal abnormal vision)

Downloaded from mcr.aacrjournals.org on July 24, 2020. © 2016 American Association for Cancer Research.

Published OnlineFirst April 13, 2016; DOI: 10.1158/1541-7786.MCR-15-0489

mediates neural-specific 3′ UTR lengthening by inhibiting usage of proximal poly(A) signals (41). Mammalian homologs of ELAV, also known as HuR, HuB, HuC, and HuD can also bind to downstream U/GU- elements of poly(A) signals and block recruitment of CSTF2. Hu proteins have no effect on poly(A) sites that do not contain U-rich sequences (42, 43). NOVA2 (neuroendocrinal ventral antigen 2) can bind to introns and 3′ UTRs of target genes regulating both splicing and polyadenylation in the brain (44). When bound close to a poly(A) signal, NOVA2 can inhibit polyadenylation at that site. However, the repressive function of NOVA2 is not effective if the poly(A) site is distant from the NOVA2-binding site; hence, polyadenylation of a distant site can be possible (44). Therefore, the position of NOVA binding may determine which poly(A) site will be selected. In conclusion, polyadenylation seems to be tightly regulated by the polyadenylation machinery proteins as well as other auxiliary proteins including RBPs. Differential polyadenylation in RBP genes themselves is another aspect of the complexity in this regulation (45).

Different Ends of the Message: Alternative Polyadenylation

Cis-regulatory elements on 3′ UTRs have been under the spotlight for quite some time due to our improved understanding of miRNA-mediated repression of gene expression. Thus, the regulatory functions of miRNAs and RBPs are of interest to further unravel the complex nature of posttranscriptional regulation, especially in cancer cells. Yet another layer of regulation is emerging for 3′ UTRs which involves controlling the length and hence the availability of *cis*-regulatory elements. Alternative polyadenylation (APA) can then be defined as the selection of alternate poly (A) signals on the pre-mRNA that leads to the generation of isoforms of various lengths depending on the location of the alternate signal (i.e., proximal or distal). These isoforms often contain different *cis*-regulatory elements, providing another layer of regulation via the 3′ UTR.

Although APA has only recently gained a more broad attention, individual cases of APA that have functional consequences have been known. For example, immunoglobulin M, androgen receptor, Collagenase 3, DNA polymerase β, Cyclin D1 and nuclear factor of activated T-cells (NFATC) are among numerous genes that that have been shown to undergo APA (reviewed in ref. 46). Moreover, the extent of genome-wide occurrences of APA was later discovered by high-throughput methods that identified 280,857 novel poly(A) sites in the human genome. These data also confirmed the 158,533 known poly(A) sites and revealed that approximately 70% of known human genes have multiple poly (A) sites in their 3′ UTRs (4, 47). Perhaps most interestingly, the use of these poly(A) signals was shown to have tissue-specific signatures (48). For example, while transcripts in the central nervous system are characterized by distal poly(A) site usage, placenta, ovaries, and blood tend to express shorter 3′ UTR isoforms by selecting more proximal poly(A) signals (49).

Further evidence of the widespread APA usage came from T lymphocyte activation where shorter 3′ UTR isoforms are associated with the induced proliferative state of these cells (50). On the contrary, a tendency toward a global 3′ UTR lengthening is observed in normal mouse embryonic development where proliferation-related genes are downregulated and differentiation/morphogenesis genes are upregulated (51). Likewise, fish eggs and zygotes, where early development is dependent on maternally derived mRNAs, express longer 3′ UTR isoforms (52). In fact, increased 3′ UTR lengths, due to APA, are so significant during the early stages of development that even individual mouse embryonic and neural stem cells have specific APA isoforms, which provides insight into functional cell-specific heterogeneity during development (53).

The unique APA patterns in different tissues and developmental stages suggest that APA has a key role in normal physiologic settings such as in embryonic stem cells (ESC). ESCs have two important characteristics: pluripotency and the ability to proliferate indefinitely, also known as self-renewal. Recently, it was suggested that APA was one of the potential mechanisms to maintain the self-renewal property of ESCs, at least in part by Fip1, a member of the CPSF complex (54). Fip1 knockout mouse ESCs lose their undifferentiated states and self-renewal abilities due to a switch from proximal to distal poly(A) site usage, which results in the 3′ UTR lengthening of a set of genes that have roles in cell fate determination. A group of these 3′ UTR–lengthened mRNAs was shown to have decreased protein expression, suggesting that Fip1 plays a role in gene silencing. Indeed, individual silencing of APA-regulated isoforms in ESCs results in impaired self-renewal ability of ESCs, which is characterized by morphologic abnormalities and aberrant expression of ESC and differentiation markers (54). To further strengthen these observations, Fip1 is also required for the generation of induced pluripotent stem cells (iPSC) from mouse embryonic fibroblasts (MEF; ref. 54). Hence, in normal physiologic settings, different lines of evidence suggest APA to be an important regulator of gene expression by generating shorter or longer mRNA isoforms.

Another interesting aspect of APA-regulated gene expression concerns the noncoding portion of the transcriptome. In addition to protein-coding transcripts, polyadenylated long noncoding RNA (lncRNAs) are also likely to go through APA. LncRNAs are transcripts that are longer than several hundred nucleotides generally with no coding potentials. While 66% of mouse lncRNA genes undergo APA (55), approximately 40% of human lncRNA genes have at least two poly(A) signals, suggesting that abundant APA produces an array of lncRNA isoforms (56). Indeed, there are examples of APA-regulated lncRNAs with different functions. For example, lncRNA NEAT1 (nuclear paraspeckle assembly transcript 1), plays a role in the formation of paraspeckles, storage sites for A-to-I edited mRNAs, which may be released into the cytoplasm under certain stress conditions (57). NEAT1 is subject to APA and only the longer isoform (NEAT1_2) is required for paraspeckle formation (58). Interestingly, one of the paraspeckle proteins, HNRNP K (heterogeneous nuclear ribonucleoprotein K), favors the production of the long NEAT1_2 isoform by preventing the binding of the CFIm complex to the vicinity of the alternative poly(A) site of NEAT1_1. Therefore, accumulation of NEAT1_2, but not NEAT1_1, along with paraspeckle proteins initiates paraspeckle construction (58).

On the basis of multiple poly(A) signals in transcribed genes, APA emerges as a factor to regulate abundance and function of lncRNA isoforms in addition to the protein-coding mRNAs. These findings suggest that APA is a common mechanism regulating the complexity of the transcriptome.

Possible Inducers of APA

The detailed mechanism of how APA is regulated in normal and in cancer cells remains to be elucidated, however; possible

Downloaded from mcr.aacrjournals.org on July 24, 2020. © 2016 American Association for Cancer Research.

Published OnlineFirst April 13, 2016; DOI: 10.1158/1541-7786.MCR-15-0489

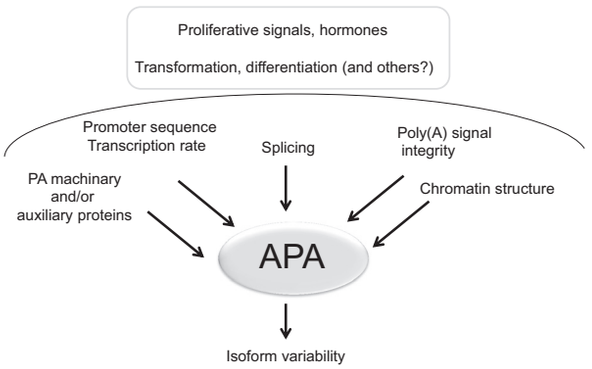


Figure 2.
Potential APA inducers. Proliferative signals, hormones, differentiation factors can induce APA. More specifically, altered expression of polyadenylation machinery proteins (and possibly other auxiliary proteins), type of promoter, rate of transcription, splicing patterns, local chromatin structure and integrity of the poly(A) signal are all possible factors to regulate APA.

mechanistic explanations are beginning to emerge (Fig. 2). One explanation is that some of the subunits of the multi-protein polyadenylation machinery have active roles in the selection of alternate signals. Perhaps one of the strongest candidates to be an APA regulator is CSTF2, a member of the cleavage stimulatory factor complex that binds downstream of the cleavage sites. Depletion of CSTF2 (and its paralog CSTF2τ) results in increased usage of distal poly(A) sites in HeLa cells, supporting a role for CSTF2 in the selection of poly(A) sites (59). Additional evidence for the selective action of CSTF2 was observed in mouse embryonic stem cells that lacked *Cstf2* expression, which resulted in the loss of differentiation potential toward the endodermal lineage, while some other differentiation pathways were still intact (60). This finding also may suggest that CSTF2 has a selective function and does not simply alter 3′ UTR lengths of all genes in all cell types.

In line with the genome-wide 3′ UTR shortening due to proliferative signals, we reported a link between signaling pathways and APA in breast cancers. E2 (estrogen) and EGF treatment in estrogen receptor–positive and EGF receptor–positive breast cancer cells, respectively, resulted in upregulation of CSTF2 and an increase in 3′ UTR shortening of all tested genes in breast cancer cells (61, 62). While E2F family of transcription factors has been implicated in CSTF2 upregulation, further studies are needed to fully uncover the mechanisms causing upregulation of CSTF2 as part of cancer-associated signaling pathways (63).

Another regulator of APA is the CFIm complex. Interestingly, depletion of CFIm25 and CFIm68 but not of CFIm59 was shown to lead to proximal polyadenylation site selection in HEK293 cells (64, 65). Similarly, CFIm25 knockdown in HeLa cells resulted in a significant shortening of 1,450 of 1453 mRNAs detected by RNA-sequencing (RNA-seq). The same study analyzed RNA-seq data from TCGA database also found 59 of 60 3′ UTR shortening events in low versus high CFIm25-expressing glioblastoma cells (66).

The CPSF complex also has been implicated in APA through Fip1/CPSF-mediated self-renewal of ESCs, which provided mechanistic insight into selective APA of target mRNAs and that only certain cell fate–related genes are regulated by APA. It turns out the distance between poly(A) signals is important for Fip1-mediated APA in ESCs. When the two poly(A) signals are distant and Fip1 activity is high (as in ESCs), proximal poly(A) signals are selected before the distal ones, even though distal signals may be more conserved and stronger. However, when 3′-end processing activity is low, cleavage and polyadenylation efficiency drops at the proximal site, and favors distal polyadenylation. On the contrary, when the two poly(A) signals are close to each other, Fip1 and CSTF compete for the limited binding space between the poly(A) signals. High Fip1 binding causes selection of the stronger distal poly(A) signal by preventing the binding of CSTF to the same region. Similarly, when Fip1 levels are low, CSTF binds to the region in between the two poly(A) signals and induces selection of the nearby proximal site (54). This level of mechanistic insight into the bimodal regulation of APA is important to understand the selectivity of APA in physiologic as well as disease settings.

In addition, as a part of the CPSF complex, CPSF73 was shown to translocate from nucleus to the cytosol under oxidative stress conditions resulting in a significant inhibition of polyadenylation activity in prostate cancers (67).

Overall, at least some polyadenylation machinery components seem to be responsible for the selection of alternate poly(A) signals. It will be interesting to reveal under which circumstances these proteins function to induce APA and what other auxiliary proteins may have roles during APA.

An additional aspect of APA regulation is dependent on the fact that polyadenylation occurs cotranscriptionally, meaning that the transcriptional machinery controlling initiation, rate of RNA polymerase, and splicing are likely to affect polyadenylation efficiency and specificity (68). In line with this possibility, novel findings have elegantly linked transcriptional initiation and RBPs to APA. APA in the developing nervous system in flies and mammals seems to favor isoforms of very long 3′ UTRs. For example, ELAV directly binds to the strong proximal polyadenylation signals of several target mRNAs to suppress short 3′ end formation (41, 69). As there is no strict motif for ELAV binding around these proximal poly(A) signal sites, the molecular explanation of how ELAV acts to block the proximal signal is still unknown. However, Oktaba and colleagues have recently shown that the native promoters of target genes are needed for ELAV recruitment to produce a longer 3′ UTR isoform. These data suggest that promoter regions with a GAGA element, known to bind to Trl (Trithorax like), tend to halt RNA Pol II and recruit ELAV to the site of the proximal poly(A) signal on the nascent transcript. Indeed, ChIP-seq analysis of *Drosophila* embryos shows an enrichment of ELAV at the promoter and the proximal poly(A) sites of target genes that tend to produce long 3′ UTR isoforms (70). Therefore, studies of the link between promoter and poly(A) site selection will be of high interest to determine the extent of this phenomena in mammalian and cancer cells, especially because ELAV homologs show similar functions during neuronal differentiation in humans (71).

On the other hand, U1 snRNP (small nuclear ribonucleoprotein), an essential component of the spliceosome, suppresses premature cleavage and polyadenylation within introns. U1 depletion leads to the activation of intronic poly(A) signals and causes genome-wide APA (reviewed in ref. 72). Furthermore, an in-depth analysis of several human tissues and cell line transcriptomes revealed a strong correlation between alternative splicing and alternative cleavage/polyadenylation patterns, suggesting coordinated regulation of these processes (73).

Downloaded from mcr.aacrjournals.org on July 24, 2020. © 2016 American Association for Cancer Research.

Published OnlineFirst April 13, 2016; DOI: 10.1158/1541-7786.MCR-15-0489

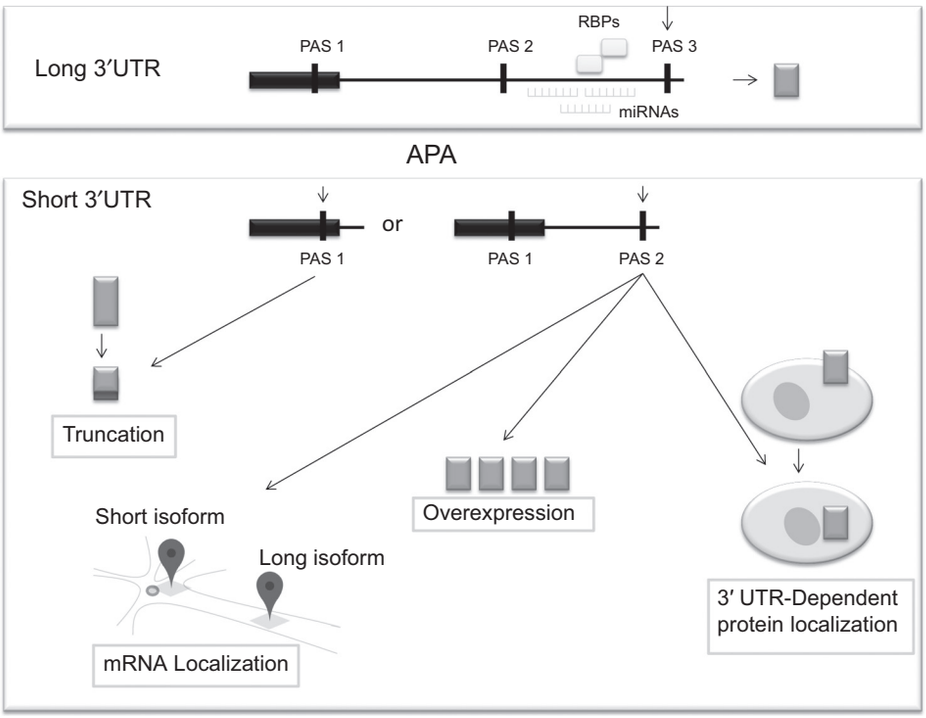


Figure 3.
Consequences of APA. Top, a hypothetical mRNA with 3 possible poly(A) signals (PASs). The longest isoform utilizes the most distal PAS (PAS 3) and harbors sites for miRs and/or RBPs. Bottom, possible mRNA isoforms in case of PAS 1 and PAS 2 selection. PAS 1 is located in the coding region and hence when selected, the resulting mRNA is truncated. In case of PAS 2 selection, coding sequence is not altered; however, resulting mRNAs can have different subcellular localization, can be translated more efficiently leading to overexpression or have different "3′ UTR-dependent protein localization."

While the position of the poly(A) signal and the cotranscriptional mechanisms regulating polyadenylation are of importance, the sequence composition of the poly(A) motif itself is also critical. There is evidence that suggests alterations in the poly (A) signal sequence motif have a functional role in diseases. For instance, a beta-globin gene that was cloned from a beta-thalassemia patient was shown to have a T to C nucleotide substitution within the AATAAA sequence motif causing an extended transcript (74). In addition, an altered poly(A) signal results in the lengthening of a well know tumor suppressor, *TP53*. A SNP (rs78378222) in the 3′ UTR of *TP53* modifies the AATAAA to AATACA and impairs termination and polyadenylation of the *TP53* transcript. This SNP was further found to be associated with prostate cancer, glioma, and colorectal adenoma, but not breast cancer (75). Later, independent case–control cohorts comprising 10,290 individuals showed rs78378222 to be robustly associated with neuroblastoma (76). Several hundred other SNPs also have been detected to create or disrupt APA signals (AATAAA and other variants) that in turn can potentially affect 3′ UTR length, miRNA regulation, and eventually mRNA levels. Intriguingly, homozygous SNPs altering poly(A) signals were bioinformatically proposed to be correlated with shortened transcript lengths, altered miRNA regulation, and ultimately altered protein levels (77).

Finally, cell signaling pathways are likely to alter genome-wide APA patterns, as part of transcriptional response. In addition to E2 and EGF, the mTOR pathway was recently linked to APA. Knockdown of mTOR in human or mouse cell lines (MEFs, HEK293, and HCV29) results in an increased number of 3′ UTR lengthening events, suggesting a conserved function of mTOR in APA regulation. Using further genetic and chemical modifications, mTOR activity was shown to be required for 3′ UTR shortening of a group of transcripts, which consequently had increased protein levels (e.g., ubiquitin-mediated proteolysis pathway members), suggesting a previously uncharacterized role for mTOR in APA. However, the molecular mechanism of how 3′ UTR lengths are modulated by mTOR activity remains elusive. While more evidence is needed for mammalian systems, APA is likely to be affected by a variety of external and internal cues, including the local chromatin structure, nucleosome positioning, DNA methylation, and histone modifications (78). Therefore, epigenetic marks in cancer cells should also be taken into consideration to better understand the extent and mechanisms of deregulated APA in cancer.

Consequences of APA; Functional Implications in Cancer

The functional significance of APA-generated isoforms and its association with the availability and usage of multiple poly(A) signals in human cancers is still unknown. To evaluate the overall effects of APA in cancer cells, we will group APA events to the following categories of alterations: (i) coding sequence; (ii) mRNA localization; (iii) protein level; and (iv) protein localization (Fig. 3).

To begin understanding the consequences of APA, it is important to consider the positions of the alternate poly(A) signals. APA and alternative splicing can occasionally operate hand in hand. In cases where proximal poly(A) signals are within introns or terminal exons, prematurely truncated proteins with potentially different and/or opposing functions can be generated. Such truncated proteins may interfere with the function of the full-length protein. One interesting example of how such a truncated isoform can alter the normal function of the wild-type protein is EPRS (glutamyl-prolyl tRNA synthetase). EPRS is a member of the GAIT (gamma-interferon-activated inhibitor of translation) complex that binds to target mRNAs such as VEGFA to inhibit their translation. Full-length EPRS mediates mRNA binding to the

Downloaded from mcr.aacrjournals.org on July 24, 2020. © 2016 American Association for Cancer Research.

Published OnlineFirst April 13, 2016; DOI: 10.1158/1541-7786.MCR-15-0489

GAIT complex; however, the truncated EPRS isoform (generated by APA) shields GAIT element–bearing target transcripts from the inhibitory GAIT complex (59). Considering the repressive effect of GAIT complex on multiple mRNA targets, consequence of this shielding could have broad importance within the context of cancer.

A more global indication of intronic poly(A) site usage is provided by the receptor tyrosine kinases (RTK). A combinatorial analysis of expressed sequence tags (EST) and RT-PCR revealed the existence of RTK mRNA isoforms that were predicted to encode dominant-negative and secreted variants. Such soluble isoforms can act as "decoy" receptors of the RTK signaling pathway with potential implications in cancer therapy (79, 80). Therefore, it is of great interest to investigate potential regulators of such intronic poly(A) signal recognition and/or suppression in cancer cells. Considering that such isoforms could go undetected at the mRNA or protein levels because we simply do not investigate their existence, it is plausible to expect more examples of proximal poly(A) signal usage that results in truncated proteins that may oppose the function of the full-length isoforms.

APA isoforms generated from the same gene may only differ in the length of their 3′ UTRs and the *cis*-regulatory elements (i.e., miR and/or RBP-binding sites) that they harbor. Thus, APA may cause increased abundance of proto-oncogenes by shortening the 3′ UTR, which effectively eliminates the negative regulatory binding sites for *trans*-factors. Relief from the negative regulators may then cause increased translation of the mRNA without transcriptional upregulation. For example, in a subset of mantle cell lymphomas, *CCND1* (Cyclin D1), 3′ UTR shortening prevents the miRNA-mediated repression and leads to an additional increase in *CCND1*, which correlates with both increased proliferation of the lymphoma cells and decreased overall survival of patients (81). Another proto-oncogene activation case was reported for Insulin-like growth factor 2 mRNA-binding protein 1 (*IGF2BP1*). Shortening of the 3′ UTR of the *IGF2BP1* transcript results in a more profound oncogenic transformation compared with the longer isoform, again showing that loss of 3′ UTR-repressive elements through APA can promote the oncogenic phenotype (82).

Finally, a link between APA and cancer also was shown for hormone-responsive APA, as mentioned above. We detected E2-dependent upregulation correlated with 3′ UTR shortening of *CDC6* (cell division cycle 6), a major regulator of DNA replication, in breast cancer cells. Thus, as a result of APA, the short 3′ UTR isoform of *CDC6* evades miRNA-mediated repression, which leads to increased CDC6 protein production and S-phase entry (61).

While increasing protein abundance is the best characterized function of APA, this view has recently been challenged. On the basis of 3′ end sequencing and quantitative mass spectrometry, 3′ UTR shortening was shown to have only limited effect on protein abundance in proliferating murine and human T cells (83). It is indeed reasonable to expect that not every APA event will correlate with higher protein levels, especially considering the existence of other posttranscriptional and posttranslational regulations. Interestingly, consequences of APA other than increased protein abundance recently have been reported.

mRNA localization appears to be an important posttranscriptional step to modulate protein levels and their functions. Spatial distribution of mRNAs provides efficient local enrichment of the protein to be translated and even assembly of multiprotein complexes in a given subcellular region of the cell (84). For example, *BDNF* (brain-derived neurotrophic factor) mRNA utilizes APA, generating short and long 3′ UTR isoforms. The short 3′ UTR isoform localizes in soma of neurons, whereas the longer isoform is found in dendrites. These isoforms have different translational rates due to their subcellular localizations and have different roles in the morphogenesis of dendritic spines (85).

Many 3′ UTRs also function as scaffolds to tether various RBPs that are likely to alter the localization and even function of the protein to be translated. This type of regulation recently has been reported for transmembrane proteins. Berkovits and Mayr coined the term "3′ UTR-dependent protein localization" based on their observations on five different proteins (CD47, CD44, ITGA1, TNFRSF13C, and TSPAN13) that are derived from HuR (also known as ELAVL1) bound mRNAs. HuR binding to uridine-rich regions of longer UTRs results in recruitment of SET (SET nuclear proto-oncogene) to the site of translation on the rER (rough endoplasmic reticulum), which allows SET to bind to the newly translated cytoplasmic domain of the protein and translocate it to the plasma membrane via activated RAC1 (ras-related C3 botulinum toxin substrate 1; ref. 86). Knockdown of HuR results in decreased surface expression of these proteins without changing the total protein or mRNA levels. The translocation of one of these proteins (CD47) to the plasma membrane is a posttranslational mechanism independent from RNA localization but dependent on the tethering properties of the longer 3′ UTR isoform. Accordingly, CD47 protein has different functions depending on whether it encoded by the short or long 3′ UTR isoform.

Although overall we currently have a limited understanding of the consequences of APA in cancer cells, emerging evidence suggests a functional role for APA-generated isoforms. Therefore, we anticipate that more examples of APA-regulated isoforms will come to light in the very near future and that each case will need to be experimentally examined for their functional potential, other than just to increase protein abundance.

Notably, APA can cause significant isoform diversity and deregulation of APA has already been implicated in diseases other than cancer. For example, FUS (fused in sarcoma), an RBP that has been implicated in neurodegeneration, affects both alternative splicing and alternative polyadenylation of thousands of neuronal mRNAs. Mutations in FUS are associated with amyotrophic lateral sclerosis (ALS), suggesting that APA and its cumulative downstream effects is one of the mechanisms involved in the pathogenesis of ALS. APA deregulation also has been implicated in myotonic dystrophy, a disease where *DMPK* (dystrophia myotonica protein kinase) and *CNBP* (CCHC-type zinc finger, nucleic acid–binding protein) genes undergo repeat expansions, CTG or CCTG, respectively. Transcribed CUG/CCUG-containing mutant RNAs aggregate as nuclear foci and sequester RBPs such as MBNL, which causes disruptions in alternative splicing and APA (reviewed in ref. 87).

Finally, disease-specific APA signatures are observed in human heart disease (88). Here, the loss of PABPN1 contributes to 3′UTR shortening of at least some of the candidate genes underlying the failing heart phenotype.

Together, these data support that APA will emerge as one of the mechanisms contributing to gene expression deregulation in numerous diseases including cancer.

Downloaded from mcr.aacrjournals.org on July 24, 2020. © 2016 American Association for Cancer Research.

Published OnlineFirst April 13, 2016; DOI: 10.1158/1541-7786.MCR-15-0489

Extent of APA in Cancer

Genetic mutations or increased copies of proto-oncogenes contribute to tumorigenesis. However, these known oncogene-activating mutations cannot fully explain the complexity of the cancer genome and transcriptome. Much of the focus in the literature has been directed towards the coding and noncoding RNAs (miRNAs, lncRNAs, etc), leaving little attention on APA and the length of the 3′ UTRs. Consequently, our current understanding of APA-based isoform variability in cancer is somewhat limited. Despite this fact, increasing evidence shows APA to be a possible regulatory mechanism of proto-oncogene activation. Following the initial findings linking 3′ UTR shortening of a set of genes and the proliferative state of cells (50), a genome-wide 3′ UTR shortening was detected in transformed cancer cell lines compared with highly proliferative nontransformed cell lines, suggesting APA has a stronger association with transformation than proliferation (82). Mouse models were also informative to understand the extent and specificity of APA in cancer. Mouse models of leukemia/lymphoma (p53 deficient) revealed 3′ UTR isoform signatures that could distinguish between (~70% accuracy) tumor subtypes with different survival characteristics. The majority of such isoforms were due to 3′ UTR shortening. Moreover, probe-based analysis of microarray data from human primary tumor samples, including breast cancer and melanoma, revealed specific APA patterns, exposing the potential of APA isoforms, as biomarkers for cancer diagnosis and prognosis (89).

Another study demonstrated that APA can be cell-type specific when compared with a nontumorigenic mammary epithelial cell line (MCF10A) and a breast cancer cell line (MCF7, ER^+) and detected a global shift toward proximal poly(A) signal usage in MCF7 cells. On the contrary, a general 3′ UTR lengthening was observed in MDA-MB-231 (ER^-) breast cancer cells (90).

To further investigate the extent of APA in cancers, 358 TCGA Pan-Cancer tumor/normal pairs were examined and found to have 1,346 genes with recurrent and tumor specific APA patterns in seven different tumor types, where approximately 90% of APA events were 3′ UTR shortening (91).

Our group has taken advantage of the APADetect tool to analyze gene expression data of more than 500 triple-negative breast cancer patients compared with normal breast tissue. These investigations revealed APA based 3′ shortening events (70%, 113 of 165) of mostly proliferation-related transcripts in patients. Representative shortening events were experimentally verified in breast cancer patient samples and cell lines. In addition, a set of the 3′ UTR shortening events were correlated with relapse free survival times, which also supports APA based proto-oncogene activation (62). The use of such meta-analysis tools of gene expression data is very beneficial for detecting APA-generated isoforms that may have been overlooked in conventional analysis. Another promise of this approach is uncovering prognostic biomarkers given the specificity of APA isoform profiles in different tumors and/or between closely related subtypes of tumors (89, 92).

How to Detect APA Isoforms

Given the vast number of poly(A) signals in the human genome, it is important to detect which signals are utilized to identify which isoforms will be generated. Here, we summarize exemplary tools to identify and study APA from both RNA-seq data and microarray gene expression data. Early efforts to identify APA-generated isoforms were mainly focused on EST databases, specifically 3′ end sequenced ESTs (93). Currently, RNA-seq–based methods are capable of analyzing expression profiles of transcripts that cover the whole 3′ UTR at single base resolution; therefore, are able to screen the whole transcriptome for APA events. While RNA-seq is more likely to provide more information on the abundance 3′ UTR isoforms and *de novo* poly(A) signal usage, potential read depletion near 3′ ends due to random priming and differential PCR amplification are potential issues (94). Recently, specialized approaches have been developed to detect APA isoforms. PolyA-Seq is a method for high-throughput sequencing of the 3′ ends of polyadenylated transcripts by synthesizing oligo(dT) primed first strand and randomer primed second strand (4). This method has the potential of providing the precise locations of poly(A) sites; however, it also suffers from internal priming artifacts (for a review, see ref. 95). Poly(A) site sequencing (PAS-Seq) uses oligo(dT) primed first strand synthesis, including a SMART adaptor (a hybrid oligonucleotide that has three guanine ribonucleotides linked to the 5′ end of a DNA linker) in the reaction. The reverse transcriptase then extends the cDNA until the 5′ end of the SMART adaptors, generating cDNAs with linkers attached to both ends, eliminating further tailing or linker ligation steps required for conventional RNA-seq (96). In 3Seq (3′-end sequencing for expression quantification), first strand is synthesized with an oligo(dT) primer that has a 3′ adapter for second strand synthesis. However, internal priming and/or polymerase slippage during oligo(dT) priming are potential risks. To address these issues, 3′ region extraction and deep sequencing (3′ READS) approach captured poly(A) tailed mRNAs using an oligo followed by adapter ligation to reveal over 5,000 previously overlooked poly(A) sites (55). The same study showed that distal polyadenylation site usage, regardless of intron or exon locations, is more abundant during embryonic development and cell differentiation. Finally, poly (A)-position profiling by sequencing (3P-Seq) method uses a splint ligation to attach adapters to mRNA (97). However, this approach may face a ligation-bias due to efficiency of multiple enzymatic reactions, affecting the overall quantification of APA events (98) (reviewed in ref. 99).

Several specialized bioinformatic tools also have been developed to investigate APA patterns using standard RNA-seq and gene expression data (Fig. 4). A list of current tools is given in Table 1.

DaPars (Dynamic analyses of Alternative PolyAdenylation from RNA-Seq; ref. 91) is a recent method that analyzes RNA-seq data and has various advantages over other tools, including 3USS (100) and MISO (101). For example, 3USS does not readily provide a comparison between groups of samples with respect to differential APA events. MISO utilizes the PolyA_DB database but may not identify novel APA events. On the other hand, DaPars performs *de novo* identification and quantification of dynamic APA events between samples regardless of any prior APA annotation. However, unlike microarray datasets found in public repositories like NCBI-GEO, RNA-seq datasets still have not been adopted at a large scale, due to higher costs and more complex data analysis. In fact, several RNA-seq repositories such as the European Genome-Phenome Archive (EGA) do not provide public access to all of the data, but govern these datasets by Data Access Committees (DACs). Hence, RNA-seq data are comparably less accessible.

Downloaded from mcr.aacrjournals.org on July 24, 2020. © 2016 American Association for Cancer Research.

Published OnlineFirst April 13, 2016; DOI: 10.1158/1541-7786.MCR-15-0489

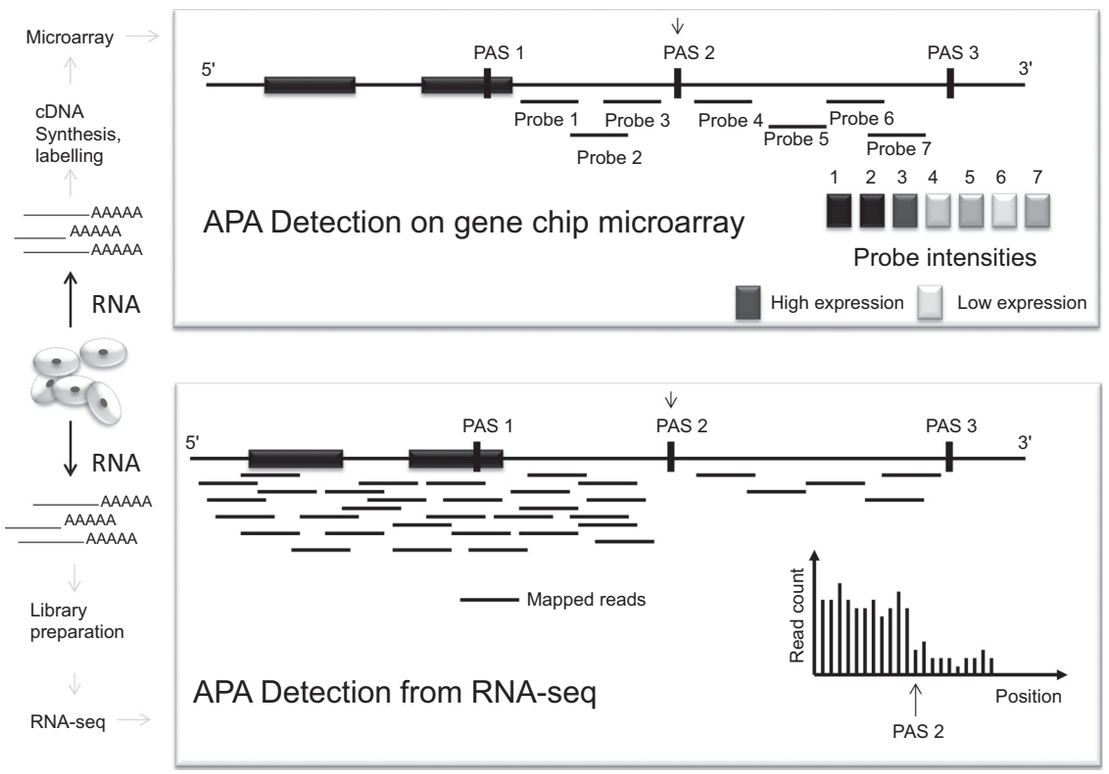


Figure 4.
Detection of APA-generated isoforms. Top, how APA events can be detected by probe level analysis of GeneChip microarrays. For an exemplary transcript, seven probes are split into two subsets; the proximal subset (probes 1, 2, and 3) and the distal subset (probes 4, 5, 6, and 7). Probe level analysis of this transcript's probe set indicates differential expression of the proximal subset compared with the distal subset. Significantly high signal intensity for the proximal subset reveals a shortening event suggesting that the alternative polyadenylation site PAS 2 is active in the analyzed sample. Bottom, APA event discovery by analyzing the coverage of mapped reads of an RNA-seq experiment. The coverage graph shows changes in the expression level at single base resolution. Drops in coverage graphs indicate shortening events suggesting utilization of alternative PASs such as PAS 2.

Fortunately, microarray data that have been collected for the primary purpose of quantifying gene expression can also be reanalyzed by computational methods retrospectively for the detection of APA events. Specifically, probe sets are generally designed from the 3′ UTRs, making it possible to identify an APA event via increased signal intensities of probes that map upstream of known poly(A) signals. The drawback of microarray-based methods is that they are limited to transcripts for which the probe set is divided into such proximal and distal sets with respect to a poly(A) site. For example, on the Affymetrix Human Genome U133 Plus 2.0 array, there are 3,683 probe sets that are divided into at least two sets by the poly(A) sites reported in the PolyA_DB database. Moreover, dependence on poly(A) site annotation databases such as the PolyA_DB can be somewhat limiting for identification of novel poly(A) sites; however, it is possible to identify candidate novel poly(A) sites by seeking break sites where probe set means differ between each other. For such probe-based analysis of microarray data, PLATA (50), APADetect (62), and ERI (102) are dependent on the PolyA_DB database for poly(A) signal position information. An important advantage of APADetect is that it can readily be used as a meta-analysis tool capable of working with multiple Affymetrix datasets from both human and mouse platforms.

Another important aspect with practical value is that the raw RNA-seq reads take about 20 gigabytes for each sample, prohibiting a quick analysis on a personal computer. Aligning raw reads to

Table 1. Available software tools for APA detection

Tool	Data source	*De novo* identification	Available as	Link	Reference
PLATA	Exon array	No	Python scripts	http://sandberg.cmb.ki.se/plata/	50
MISO	RNA-seq	No	Python package	http://genes.mit.edu/burgelab/miso/	101
ERI	GeneChip microarray	No	R program	Software Supplement at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0031129	102
DaPars	RNA-seq	Yes	Python program	https://github.com/ZhengXia/DaPars	91
3USS	RNA-seq	Yes	Web server	http://circe.med.uniroma1.it/3uss_server/	100
APADetect	GeneChip microarray	No	Java program	http://www.ceng.metu.edu.tr/~tcan/APADetect/	61, 62

Downloaded from mcr.aacrjournals.org on July 24, 2020. © 2016 American Association for Cancer Research.

Published OnlineFirst April 13, 2016; DOI: 10.1158/1541-7786.MCR-15-0489

the human reference genome takes approximately 1.5 days per sample. In contrast, raw microarray CEL files for single samples are around 30 megabytes and analysis using microarray data takes a couple of minutes to complete by probe based tools such as APADetect. In conclusion, our efforts to compare RNA-seq and microarray data demonstrate significant running time and space costs associated with RNA-Seq analysis to be a major obstacle in rapid and practical use of these datasets for detection of APA events. While probe-based tools can easily be used for meta-analysis of existing data, this approach may not detect certain APA events due to the limitation of original probe position designs. Therefore, researchers must weigh the abovementioned constraints for their specific questions.

Conclusion

Strong evidence shows that APA is a general mechanism of gene expression regulation during different physiologic states. Deregulated APA has several downstream consequences that may affect protein function(s) that are dependent or independent of mRNA and/or protein levels. Therefore, a deep understanding of APA is likely to yield novel cancer genes that may have been overlooked with conventional gene expression analysis approaches, simply because we do not look for such isoforms. While RNA-seq is overtaking other transcriptome analysis approaches, we strongly urge researchers to consider the benefit of reanalyzing the vast amount of existing microarray data for APA detection.

Overall, APA along with alternative splicing adds another layer to the complexity of gene expression regulation. Unraveling this complexity may yield new information to cancer researchers. One interesting aspect of this new information is the potential therapeutic applications. For example, given its role in suppressing intronic poly(A) signals, blocking U1 may enhance intronic polyadenylation in numerous genes, some of which encode receptor tyrosine kinases (RTK), thereby producing soluble variant RTK isoforms lacking the transmembrane domains. These soluble variants can then act in a dominant-negative way to block receptor signaling. In addition, U1 silencing can also promote the use of intronic poly(A) signals in other genes such as *VEGFR2* pre-mRNA, which could result in elevated production of the soluble decoy VEGFR2 protein and attenuation of angiogenic signals (80). More examples of such approaches are likely to emerge as the extent of APA in cancer cells becomes clear.

In conclusion, APA represents a molecular explanation for protein activation and/or inactivation in complex cancer cases of unknown cause. While the extent and consequences of all APA events are not known in noncancerous and cancer cells, APA events, individually and/or cumulatively, are likely to contribute to the initiation and/or maintenance of tumors. Therefore, a deeper understanding of APA promises to yield new information on gene expression regulation and hopefully will pave the way to improve our understanding of cancer and help in the development of novel diagnostic and therapeutic approaches.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

The authors thank Dr. R. Alisch and Dr. M. Muyan for critical reading of the manuscript.

Grant Support

The APA work is funded by grants TUBITAK 112S478 and 114Z884.

Received December 16, 2015; revised March 30, 2016; accepted March 30, 2016; published OnlineFirst April 13, 2016.

References

1. Marzluff WF, Wagner EJ, Duronio RJ. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. Nat Rev Genet 2008;9:843–54.
2. Proudfoot N. Poly(A) signals. Cell 1991;64:671–4.
3. Colgan DF, Manley JL. Mechanism and regulation of mRNA polyadenylation. Genes Dev 1997;11:2755–66.
4. Derti A, Garrett-Engele P, Macisaac KD, Stevens RC, Sriram S, Chen R, et al. A quantitative atlas of polyadenylation in five mammals. Genome Res 2012;22:1173–83.
5. Millevoi S, Vagner S. Molecular mechanisms of eukaryotic pre-mRNA 3′ end processing regulation. Nucleic Acids Res 2010;38:2757–74.
6. Tian B, Graber JH. Signals for pre-mRNA cleavage and polyadenylation. Wiley Interdiscip Rev RNA 2012;3:385–96.
7. Zarudnaya MI, Kolomiets IM, Potyahaylo AL, Hovorun DM. Downstream elements of mammalian pre-mRNA polyadenylation signals: primary, secondary and higher-order structures. Nucleic Acids Res 2003;31:1375–86.
8. Kuhn RM, Karolchik D, Zweig AS, Wang T, Smith KE, Rosenbloom KR, et al. The UCSC Genome Browser Database: update 2009. Nucleic Acids Res 2009;37:D755–61.
9. Kaufmann I, Martin G, Friedlein A, Langen H, Keller W. Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly (A) polymerase. EMBO J 2004;23:616–26.
10. Mandel CR, Bai Y, Tong L. Protein factors in pre-mRNA 3′-end processing. Cell Mol Life Sci 2008;65:1099–122.
11. Chan SL, Huppertz I, Yao C, Weng L, Moresco JJ, Yates JR, et al. CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3′ processing. Genes Dev 2014;28:2370–80.
12. Schönemann L, Kühn U, Martin G, Schäfer P, Gruber AR, Keller W, et al. Reconstitution of CPSF active in polyadenylation: recognition of the polyadenylation signal by WDR33. Genes Dev 2014;28:2381–93.
13. Mandel CR, Kaneko S, Zhang H, Gebauer D, Vethantham V, Manley JL, et al. Polyadenylation factor CPSF-73 is the pre-mRNA 3′-end-processing endonuclease. Nature 2006;444:953–6.
14. Dantonel JC, Murthy KG, Manley JL, Tora L. Transcription factor TFIID recruits factor CPSF for formation of 3′ end of mRNA. Nature 1997;389:399–402.
15. Nag A, Narsinh K, Martinson HG. The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. Nat Struct Mol Biol 2007;14:662–9.
16. Glover-Cutter K, Kim S, Espinosa J, Bentley DL. RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. Nat Struct Mol Biol 2008;15:71–8.
17. Kwon C, Tak H, Rho M, Chang HR, Kim YH, Kim KT, et al. Detection of PIWI and piRNAs in the mitochondria of mammalian cancer cells. Biochem Biophys Res Commun 2014;446:218–23.
18. MacDonald CC, Wilusz J, Shenk T. The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. Mol Cell Biol 1994;14:6647–54.
19. Wilusz J, Shenk T. A uridylate tract mediates efficient heterogeneous nuclear ribonucleoprotein C protein-RNA cross-linking and functionally substitutes for the downstream element of the polyadenylation signal. Mol Cell Biol 1990;10:6397–407.
20. Takagaki Y, Manley JL. Complex protein interactions within the human polyadenylation machinery identify a novel component. Mol Cell Biol 2000;20:1515–25.

Downloaded from mcr.aacrjournals.org on July 24, 2020. © 2016 American Association for Cancer Research.

Published OnlineFirst April 13, 2016; DOI: 10.1158/1541-7786.MCR-15-0489

21. Yao Y, Song L, Katz Y, Galili G. Cloning and characterization of Arabidopsis homologues of the animal CstF complex that regulates 3′ mRNA cleavage and polyadenylation. J Exp Bot 2002;53:2277–8.
22. Cevher MA, Kleiman FE. Connections between 3′-end processing and DNA damage response. Wiley Interdiscip Rev RNA 2010;1:193–9.
23. Brown KM, Gilmartin GM. A mechanism for the regulation of pre-mRNA 3′ processing by human cleavage factor Im. Mol Cell 2003;12: 1467–76.
24. Venkataraman K, Brown KM, Gilmartin GM. Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. Genes Dev 2005;19:1315–27.
25. Coseno M, Martin G, Berger C, Gilmartin G, Keller W, Doublié S. Crystal structure of the 25 kDa subunit of human cleavage factor Im. Nucleic Acids Res 2008;36:3474–83.
26. Yang Q, Gilmartin GM, Doublié S. Structural basis of UGUA recognition by the Nudix protein CFI(m)25 and implications for a regulatory role in mRNA 3′ processing. Proc Natl Acad Sci U S A 2010;107: 10062–7.
27. Li H, Tong S, Li X, Shi H, Ying Z, Gao Y, et al. Structural basis of pre-mRNA recognition by the human cleavage factor Im complex. Cell Res 2011; 21:1039–51.
28. Yang Q, Doublié S. Structural biology of poly(A) site definition. Wiley Interdiscip Rev RNA 2011;2:732–47.
29. Birse CE, Minvielle-Sebastia L, Lee BA, Keller W, Proudfoot NJ. Coupling termination of transcription to messenger RNA maturation in yeast. Science 1998;280:298–301.
30. Proudfoot NJ. Ending the message: poly(A) signals then and now. Genes Dev 2011;25:1770–82.
31. Scorilas A. Polyadenylate polymerase (PAP) and 3′ end pre-mRNA processing: function, assays, and association with disease. Crit Rev Clin Lab Sci 2002;39:193–224.
32. Eckmann CR, Rammelt C, Wahle E. Control of poly(A) tail length. Wiley Interdiscip Rev RNA 2011;2:348–61.
33. Lutz CS, Moreira A. Alternative mRNA polyadenylation in eukaryotes: an effective regulator of gene expression. Wiley Interdiscip Rev RNA 2011; 2:23–31.
34. Subtelny AO, Eichhorn SW, Chen GR, Sive H, Bartel DP. Poly(A)-tail profiling reveals an embryonic switch in translational control. Nature 2014;508:66–71.
35. Wang L, Eckmann CR, Kadyk LC, Wickens M, Kimble J. A regulatory cytoplasmic poly(A) polymerase in Caenorhabditis elegans. Nature 2002;419:312–6.
36. Saitoh S, Chabes A, McDonald WH, Thelander L, Yates JR, Russell P. Cid13 is a cytoplasmic poly(A) polymerase that regulates ribonucleotide reductase mRNA. Cell 2002;109:563–73.
37. Rouhana L, Wang L, Buter N, Kwak JE, Schiltz CA, Gonzalez T, et al. Vertebrate GLD2 poly(A) polymerases in the germline and the brain. RNA 2005;11:1117–30.
38. Charlesworth A, Meijer HA, de Moor CH. Specificity factors in cytoplasmic polyadenylation. Wiley Interdiscip Rev RNA 2013;4:437–61.
39. Bava FA, Eliscovich C, Ferreira PG, Miñana B, Ben-Dov C, Guigó R, et al. CPEB1 coordinates alternative 3′-UTR formation with translational regulation. Nature 2013;495:121–5.
40. Batra R, Charizanis K, Manchanda M, Mohan A, Li M, Finn DJ, et al. Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. Mol Cell 2014;56:311–22.
41. Hilgers V, Lemke SB, Levine M. ELAV mediates 3′ UTR extension in the Drosophila nervous system. Genes Dev 2012;26:2259–64.
42. Dai W, Zhang G, Makeyev EV. RNA-binding protein HuR autoregulates its expression by promoting alternative polyadenylation site usage. Nucleic Acids Res 2012;40:787–800.
43. Zhu H, Zhou HL, Hasman RA, Lou H. Hu proteins regulate polyadenylation by blocking sites containing U-rich sequences. J Biol Chem 2007; 282:2203–10.
44. Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 2008;456:464–9.
45. Hu W, Liu Y, Yan J. Microarray meta-analysis of RNA-binding protein functions in alternative polyadenylation. PLoS One 2014;9:e90774.
46. Akman HB, Erson-Bensan AE. Alternative polyadenylation and its impact on cellular processes. Microrna 2014;3:2–9.
47. Tian B, Hu J, Zhang H, Lutz CS. A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res 2005; 33:201–12.
48. Ni T, Yang Y, Hafez D, Yang W, Kiesewetter K, Wakabayashi Y, et al. Distinct polyadenylation landscapes of diverse human tissues revealed by a modified PA-seq strategy. BMC Genomics 2013;14:615.
49. Zhang H, Lee JY, Tian B. Biased alternative polyadenylation in human tissues. Genome Biol 2005;6:R100.
50. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB. Proliferating cells express mRNAs with shortened 3′ untranslated regions and fewer microRNA target sites. Science 2008;320:1643–7.
51. Ji Z, Lee JY, Pan Z, Jiang B, Tian B. Progressive lengthening of 3′ untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. Proc Natl Acad Sci U S A 2009;106: 7028–33.
52. Kleppe L, Edvardsen RB, Kuhl H, Malde K, Furmanek T, Drivenes O, et al. Maternal 3′UTRs: from egg to onset of zygotic transcription in Atlantic cod. BMC Genomics 2012;13:443.
53. Velten L, Anders S, Pekowska A, Järvelin AI, Huber W, Pelechano V, et al. Single-cell polyadenylation site mapping reveals 3′ isoform choice variability. Mol Syst Biol 2015;11:812.
54. Lackford B, Yao C, Charles GM, Weng L, Zheng X, Choi EA, et al. Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. EMBO J 2014;33:878–89.
55. Hoque M, Ji Z, Zheng D, Luo W, Li W, You B, et al. Analysis of alternative cleavage and polyadenylation by 3′ region extraction and deep sequencing. Nat Methods 2013;10:133–9.
56. Zhang S, Han J, Zhong D, Liu R, Zheng J. Genome-wide identification and predictive modeling of lincRNAs polyadenylation in cancer genome. Comput Biol Chem 2014;52:1–8.
57. Prasanth KV, Prasanth SG, Xuan Z, Hearn S, Freier SM, Bennett CF, et al. Regulating gene expression through RNA nuclear retention. Cell 2005; 123:249–63.
58. Naganuma T, Nakagawa S, Tanigawa A, Sasaki YF, Goshima N, Hirose T. Alternative 3′-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. EMBO J 2012;31:4020–34.
59. Yao CG, Biesinger J, Wan J, Weng LJ, Xing Y, Xie XH, et al. Transcriptome-wide analyses of CstF64-RNA interactions in global regulation of mRNA alternative polyadenylation. Proc Natl Acad Sci U S A 2012;109:18773.
60. Youngblood BA, Grozdanov PN, MacDonald CC. CstF-64 supports pluripotency and regulates cell cycle progression in embryonic stem cells through histone 3′ end processing. Nucleic Acids Res 2014;42: 8330–42.
61. Akman BH, Can T, Erson-Bensan AE. Estrogen-induced upregulation and 3′-UTR shortening of CDC6. Nucleic Acids Res 2012;40:10679–88.
62. Akman HB, Oyken M, Tuncer T, Can T, Erson-Bensan AE. 3′UTR shortening and EGF signaling: implications for breast cancer. Hum Mol Genet 2015;24:6910–20.
63. Elkon R, Drost J, van Haaften G, Jenal M, Schrier M, Oude Vrielink JA, et al. E2F mediates enhanced alternative polyadenylation in proliferation. Genome Biol 2012;13:R59.
64. Gruber AR, Martin G, Keller W, Zavolan M. Cleavage factor Im is a key regulator of 3′ UTR length. RNA Biol 2012;9:1405–12.
65. Martin G, Gruber AR, Keller W, Zavolan M. Genome-wide analysis of pre-mRNA 3′ end processing reveals a decisive role of human cleavage factor I in the regulation of 3′ UTR length. Cell Rep 2012;1:753–63.
66. Masamha CP, Xia Z, Yang J, Albrecht TR, Li M, Shyu AB, et al. CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. Nature 2014;510:412–6.
67. Zhu ZH, Yu YP, Shi YK, Nelson JB, Luo JH. CSR1 induces cell death through inactivation of CPSF3. Oncogene 2009;28:41–51.
68. Proudfoot NJ, Furger A, Dye MJ. Integrating mRNA processing with transcription. Cell 2002;108:501–12.
69. Elkon R, Ugalde AP, Agami R. Alternative cleavage and polyadenylation: extent, regulation and function. Nat Rev Genet 2013;14:496–506.
70. Oktaba K, Zhang W, Lotz TS, Jun DJ, Lemke SB, Ng SP, et al. ELAV links paused Pol II to alternative polyadenylation in the Drosophila nervous system. Mol Cell 2015;57:341–8.
71. Mansfield KD, Keene JD. Neuron-specific ELAV/Hu proteins suppress HuR mRNA during neuronal differentiation by alternative polyadenylation. Nucleic Acids Res 2012;40:2734–46.

Downloaded from mcr.aacrjournals.org on July 24, 2020. © 2016 American Association for Cancer Research.

Published OnlineFirst April 13, 2016; DOI: 10.1158/1541-7786.MCR-15-0489

72. Spraggon L, Cartegni L. U1 snRNP-dependent suppression of polyadenylation: physiological role and therapeutic opportunities in cancer. Int J Cell Biol 2013;2013:846510.
73. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, et al. Alternative isoform regulation in human tissue transcriptomes. Nature 2008;456:470–6.
74. Orkin SH, Cheng TC, Antonarakis SE, Kazazian HH. Thalassemia due to a mutation in the cleavage-polyadenylation signal of the human beta-globin gene. EMBO J 1985;4:453–6.
75. Stacey SN, Sulem P, Jonasdottir A, Masson G, Gudmundsson J, Gudbjartsson DF, et al. A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. Nat Genet 2011;43:1098–103.
76. Diskin SJ, Capasso M, Diamond M, Oldridge DA, Conkrite K, Bosse KR, et al. Rare variants in TP53 and susceptibility to neuroblastoma. J Natl Cancer Inst 2014;106:dju047.
77. Thomas LF, Sætrom P. Single nucleotide polymorphisms can create alternative polyadenylation signals and affect gene expression through loss of microRNA-regulation. PLoS Comput Biol 2012;8:e1002621.
78. Huang H, Liu H, Sun X. Nucleosome distribution near the 3′ ends of genes in the human genome. Biosci Biotechnol Biochem 2013;77:2051–5.
79. Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. Cell 2010;141:1117–34.
80. Vorlová S, Rocco G, Lefave CV, Jodelka FM, Hess K, Hastings ML, et al. Induction of antagonistic soluble decoy receptor tyrosine kinases by intronic polyA activation. Mol Cell 2011;43:927–39.
81. Rosenwald A, Wright G, Wiestner A, Chan WC, Connors JM, Campo E, et al. The proliferation gene expression signature is a quantitative integrator of oncogenic events that predicts survival in mantle cell lymphoma. Cancer Cell 2003;3:185–97.
82. Mayr C, Bartel DP. Widespread shortening of 3′UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell 2009;138:673–84.
83. Gruber AR, Martin G, Müller P, Schmidt A, Gruber AJ, Gumienny R, et al. Global 3′ UTR shortening has a limited effect on protein abundance in proliferating T cells. Nat Commun 2014;5:5465.
84. Mingle LA, Okuhama NN, Shi J, Singer RH, Condeelis J, Liu G. Localization of all seven messenger RNAs for the actin-polymerization nucleator Arp2/3 complex in the protrusions of fibroblasts. J Cell Sci 2005;118:2425–33.
85. Lau AG, Irier HA, Gu J, Tian D, Ku L, Liu G, et al. Distinct 3′UTRs differentially regulate activity-dependent translation of brain-derived neurotrophic factor (BDNF). Proc Natl Acad Sci U S A 2010;107:15945–50.
86. Berkovits BD, Mayr C. Alternative 3′ UTRs act as scaffolds to regulate membrane protein localization. Nature 2015;522:363–7.
87. Meola G, Cardani R. Myotonic dystrophies: An update on clinical aspects, genetic, pathology, and molecular pathomechanisms. Biochim Biophys Acta 2015;1852:594–606.
88. Creemers EE, Bawazeer A, Ugalde AP, van Deutekom HW, van der Made I, de Groot NE, et al. Genome-wide polyadenylation maps reveal dynamic mRNA 3′-end formation in the failing human heart. Circ Res 2016;118:433–8.
89. Singh P, Alley TL, Wright SM, Kamdar S, Schott W, Wilpan RY, et al. Global changes in processing of mRNA 3′ untranslated regions characterize clinically distinct cancer subtypes. Cancer Res 2009;69:9422–30.
90. Fu Y, Sun Y, Li Y, Li J, Rao X, Chen C, et al. Differential genome-wide profiling of tandem 3′ UTRs among human breast cancer and normal cells by high-throughput sequencing. Genome Res 2011;21:741–7.
91. Xia Z, Donehower LA, Cooper TA, Neilson JR, Wheeler DA, Wagner EJ, et al. Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3′-UTR landscape across seven tumour types. Nat Commun 2014;5:5274.
92. Morris AR, Bos A, Diosdado B, Rooijers K, Elkon R, Bolijn AS, et al. Alternative cleavage and polyadenylation during colorectal cancer development. Clin Cancer Res 2012;18:5256–66.
93. Gautheret D, Poirot O, Lopez F, Audic S, Claverie JM. Alternate polyadenylation in human mRNAs: a large-scale analysis by EST clustering. Genome Res 1998;8:524–30.
94. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009;10:57–63.
95. Sun Y, Fu Y, Li Y, Xu A. Genome-wide alternative polyadenylation in animals: insights from high-throughput technologies. J Mol Cell Biol 2012;4:352–61.
96. Shepard PJ, Choi EA, Lu J, Flanagan LA, Hertel KJ, Shi Y. Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. RNA 2011;17:761–72.
97. Jan CH, Friedman RC, Ruby JG, Bartel DP. Formation, regulation and evolution of Caenorhabditis elegans 3′UTRs. Nature 2011;469:97–101.
98. Hafner M, Renwick N, Brown M, Mihailović A, Holoch D, Lin C, et al. RNA-ligase-dependent biases in miRNA representation in deep-sequenced small RNA cDNA libraries. RNA 2011;17:1697–712.
99. Nussbacher JK, Batra R, Lagier-Tourenne C, Yeo GW. RNA-binding proteins in neurodegeneration: Seq and you shall receive. Trends Neurosci 2015;38:226–36.
100. Le Pera L, Mazzapioda M, Tramontano A. 3USS: a web server for detecting alternative 3′UTRs from RNA-seq experiments. Bioinformatics 2015;31:1845–7.
101. Katz Y, Wang ET, Airoldi EM, Burge CB. Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat Methods 2010;7:1009–15.
102. Lembo A, Di Cunto F, Provero P. Shortening of 3′UTRs correlates with poor prognosis in breast and lung cancer. PLoS One 2012;7:e31129.

Downloaded from mcr.aacrjournals.org on July 24, 2020. © 2016 American Association for Cancer Research.

Published OnlineFirst April 13, 2016; DOI: 10.1158/1541-7786.MCR-15-0489

AACR American Association for Cancer Research

Molecular Cancer Research

Alternative Polyadenylation: Another Foe in Cancer

Ayse Elif Erson-Bensan and Tolga Can

Mol Cancer Res 2016;14:507-517. Published OnlineFirst April 13, 2016.

Updated version Access the most recent version of this article at: doi:10.1158/1541-7786.MCR-15-0489

Cited articles This article cites 102 articles, 30 of which you can access for free at: http://mcr.aacrjournals.org/content/14/6/507.full#ref-list-1

Citing articles This article has been cited by 6 HighWire-hosted articles. Access the articles at: http://mcr.aacrjournals.org/content/14/6/507.full#related-urls

E-mail alerts Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, use this link http://mcr.aacrjournals.org/content/14/6/507. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.

Downloaded from mcr.aacrjournals.org on July 24, 2020. © 2016 American Association for Cancer Research.