


Review

Implications of Poly(A) Tail Processing in Repeat Expansion Diseases

Paweł Joachimiak, Adam Ciesiołka, Grzegorz Figura and Agnieszka Fiszer * 

Department of Medical Biotechnology, Institute of Bioorganic Chemistry Polish Academy of Sciences, 61-704 Poznań, Poland; pjoachimiak@ibch.poznan.pl (P.J.); a.k.ciesiolka@gmail.com (A.C.); grzegorz.figura5@gmail.com (G.F.)

* Correspondence: agnieszka.fiszer@ibch.poznan.pl

Abstract: Repeat expansion diseases are a group of more than 40 disorders that affect mainly the nervous and/or muscular system and include myotonic dystrophies, Huntington's disease, and fragile X syndrome. The mutation-driven expanded repeat tract occurs in specific genes and is composed of tri- to dodeca-nucleotide-long units. Mutant mRNA is a pathogenic factor or important contributor to the disease and has great potential as a therapeutic target. Although repeat expansion diseases are quite well known, there are limited studies concerning polyadenylation events for implicated transcripts that could have profound effects on transcript stability, localization, and translation efficiency. In this review, we briefly present polyadenylation and alternative polyadenylation (APA) mechanisms and discuss their role in the pathogenesis of selected diseases. We also discuss several methods for poly(A) tail measurement (both transcript-specific and transcriptome-wide analyses) and APA site identification—the further development and use of which may contribute to a better understanding of the correlation between APA events and repeat expansion diseases. Finally, we point out some future perspectives on the research into repeat expansion diseases, as well as APA studies.

Keywords: alternative polyadenylation; repeat expansion diseases; polyglutamine diseases; Huntington's disease; poly(A) tail



Citation: Joachimiak, P.; Ciesiołka, A.; Figura, G.; Fiszer, A. Implications of Poly(A) Tail Processing in Repeat Expansion Diseases. *Cells* **2022**, *11*, 677. <https://doi.org/10.3390/cells11040677>

Academic Editor: Adrian Israelson

Received: 17 January 2022

Accepted: 13 February 2022

Published: 15 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The main role of mRNAs is to transfer genetic information from DNA into proteins; however, mRNA also functions to provide an additional level of gene expression regulation in cells. Newly generated transcripts—before becoming translation templates—undergo modifications and processing, including splicing and the site selection of the transcript's 3' end by the generation of a poly(A) tail [1]. As a result of the latter, a single mature mRNA has a defined 3'-UTR region and a specific poly(A) tail length. Nevertheless, the overall pool of generated mRNAs contains molecules with different 3' ends. These features differ for any given mRNA—especially in the developmental stages and across various cell types—which affects its stability and overall transcript functioning.

There is a great need for the precise, molecular characterization of transcript variants to understand cell-type-specific processes and their disruption in cases of gene mutations. This includes aspects of poly(A) tail processing, which is especially crucial in neuronal cells, as features of the 3' end of an mRNA play an important role in regulating the dynamics of RNA metabolism and translation [2].

In this review, we discuss poly(A) tail processing in a context of a group of neurodegenerative and neuromuscular diseases caused by genes mutations consisting of the expansion of a sequence of repeats, e.g., CAG or CUG. A better understanding of a poly(A) tail processing in the context of neuromuscular pathology may provide valuable insight into molecular pathomechanisms triggered by transcript misprocessing and help design RNA-targeted therapies.

We briefly present some general information regarding the processes associated with polyadenylation and the methods used for their analysis. We also discuss previously collected information concerning poly(A) tail processing in mRNAs implicated in repeat expansion diseases, as well as the identified disruptions in this process. This includes the analysis of alternative polyadenylation sites for a group of transcripts containing expanded CAG repeat tracts implicated in polyglutamine (polyQ) diseases.

2. Polyadenylation in Animal Cells

2.1. Polyadenylation Process

During polyadenylation, a polyadenosine sequence—namely, a poly(A) tail—is added to the 3′ end of a transcript. Together with the removal of introns and the addition of a 5′ cap, polyadenylation constitutes a major step in pre-mRNA maturation [1,3,4]. The polyadenylation process can be divided into two major steps: first, newly transcribed pre-mRNA is cleaved and its 3′ end is generated; then, a specific enzyme—poly(A) polymerase (PAP)—generates the poly(A) tail independently from the template, starting from the cleavage site [5–7]. These two steps depend on the interplay between various *cis*- and *trans*-acting factors (Figure 1). *Cis* elements, which are typically located in 3′-UTRs, determine the position of cleavage and polyadenylation site in mRNA [8]. The most crucial *cis*-element is an RNA sequence motif called the polyadenylation signal (PAS). In animals, the canonical PAS is the hexanucleotide sequence AAUAAA [9–11]. The cleavage site is located around 10 to 30 nucleotides downstream from the PAS, between the PAS and a core downstream sequence element (DSE)—a U-/GU-rich sequence, which is itself located 14–70 nucleotides downstream of the cleavage site (Figure 1) [12]. The cleavage itself occurs just before an adenosine residue, mostly after cytosine. Another *cis*-element that is involved in the polyadenylation process is a U-rich/UGUA upstream sequence element (USE) located upstream of PAS. All three *cis*-elements are recognized by their respective *trans*-acting factors, which interact with each other through the carboxyl-terminal domain (CTD) of RNA polymerase II (RNAPII) [5]. PAS recruits cleavage and polyadenylation specificity factor (CPSF), DSE is recognized by cleavage stimulation factor (CstF), and USE recruits cleavage factor I_m (CFI_m). Cooperation between those factors leads to cleavage. The subsequent cooperation between nuclear poly(A)-binding protein (PABPN1) and PAP allows for the generation of the poly(A) tail at the cleavage site: PABPN1 acts as a type of ‘ruler’, which is important for the synthesis of an appropriately sized poly(A) tail, while PAP performs a non-templated addition of adenosine residues [13].

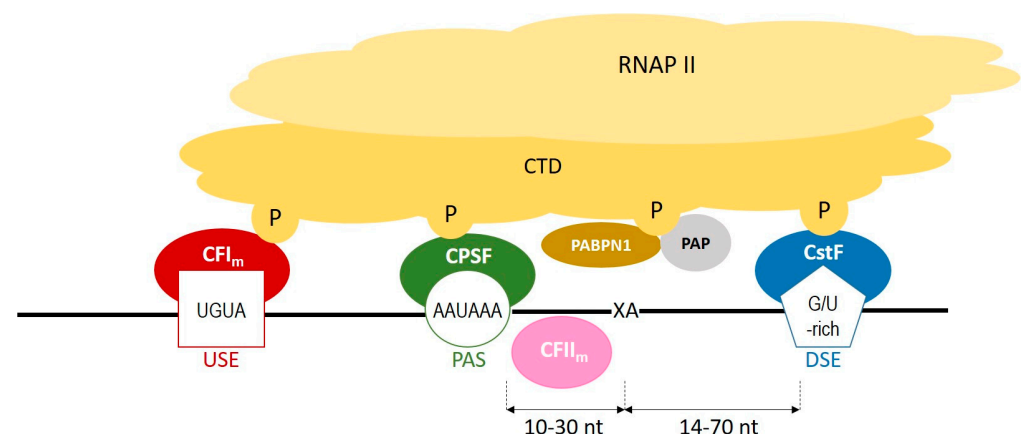


Figure 1. A scheme representing the cooperation between *cis*-elements and *trans*-factors, which are involved in a polyadenylation process. CFI_m—cleavage factor I_m; CFI_{II,m}—cleavage factor II_m; CPSF—cleavage and polyadenylation specificity factor; CstF—cleavage stimulation factor; CTD—carboxyterminal domain of RNA polymerase II; DSE—downstream element; P—phosphorylated serines of CTD; PABPN1—nuclear poly(A) binding protein; PAP—poly(A) polymerase; PAS—poly(A) signal; RNAPII—RNA polymerase II; USE—upstream element; XA—cleavage site.

2.2. Polyadenylation Roles

The poly(A) tail plays an important role in a transcript's life cycle [2]. For example, it is responsible for mediating the transport of mature mRNA into the cytoplasm through the NXF1-dependent pathway using nuclear-pore complexes embedded in the nuclear envelope [14]. Poly(A) tails are also involved in maintaining the stability of transcripts. They participate in the regulation of mRNA degradation through the process of deadenylation, which is commonly activated in post-transcriptional regulation via the sequence-specific binding of microRNA (miRNA), together with miRNA-induced silencing complex (miRISC), to the 3'-UTR [15,16]. Poly(A) tails are also important for translation regulation [17,18]. It is suggested that they interact functionally and physically with the 5' cap of mRNA, creating a closed-loop structure that promotes the initiation of translation.

In metazoa, almost all mRNAs undergo polyadenylation. One of the exceptions is replication-dependent histone protein mRNAs, which form a highly conserved stem-loop structure [19,20]. Additionally, according to a recent comprehensive study [21], nearly 50% of long non-coding RNAs (lncRNAs) undergo polyadenylation, and the resulting poly(A) tails are important for their regulation at the cellular level [22].

2.3. Length and Composition of Tails

The lengths of poly(A) tails can vary between transcripts. For example, mRNAs of highly expressed genes, such as housekeeping genes, usually possess shorter poly(A) tails, whereas poorly translated transcripts and lncRNAs have longer tails [23,24]. Generally, the length range for specifying a short or long tail is dependent on the population of transcripts analyzed. For example, in a *Caenorhabditis elegans* study, half of the transcripts possessed poly(A) tails in the range of 70 to 94 A residues; therefore, short tails were defined as those with ≤ 70 A residues, and long ones, as those with >94 A residues [23]. The existence of shorter tails may be also explained by the fact that some genes (e.g., albumin and transferrin) contain poly(A)-limiting element (PLE), which tends to restrict the initial length of poly(A) tails from the pre-mRNA stage [25,26]. Additionally, a cytoplasmic polyadenylation element (CPE) can modulate the length of a poly(A) tail after the export of the transcript from the nucleus [27]. It is located in the 3'-UTR, near PAS, and its most common sequence is UUUUAU [28]. As investigated in neurons and during early development, CPE can influence the length of the poly(A) tail through the binding of proteins to this motif (e.g., CPEB) [29]. CPEB can promote either cytoplasmic polyadenylation or deadenylation. When it is deactivated (dephosphorylated), it recruits poly(A) ribonuclease deadenylase (PARN) to deadenylate and repress the expression of mRNA. On the other hand, when it is phosphorylated, it promotes the expulsion of PARN, thereby leading to polyadenylation by germ-line development factor 2 (GLD2) [26,30]. This cytoplasmic polyadenylation process occurs in mRNAs that already contain a short poly(A) tail and usually activates translation, leading to increased protein expression.

An interesting fact is that poly(A) tails are not limited to possessing A residues: reports indicate that cytosines, guanosines, and uridines can also be incorporated into the poly(A) tail. While the role of cytosines added to the poly(A) tail remains to be elucidated, the guanylation and uridylation of poly(A) tails are quite well understood. Guanylation occurs only for longer poly(A) tails and can slow down deadenylation, delaying transcript decay [31]. Guanylation frequently occurs for long-lived transcripts with a slow turnover, e.g., transcripts encoding secreted proteins [31]. On the other hand, uridylation is usually found in short tails and marks transcripts for decay [32,33]. This process tends to be a key factor in germline development, differentiation, and early embryogenesis, where short-lived transcripts, with relatively fast turnover, predominate [33]. The lengths of poly(A) tails might also depend on circadian rhythms and the cell cycle [34–36]. By studying multiple mouse liver mRNAs, researchers demonstrated that rhythmic changes in poly(A) tail lengths were under the control of the circadian clock. Even more importantly, they presented data indicating that rhythmic poly(A) tail lengths are correlated with rhythmic protein expression [34]. With regard to the cell cycle, TAIL-Seq analysis suggests that

global RNA decay takes place during the S phase through the accumulation of terminal uridylation. On the other hand, the accumulation of terminal guanylation occurs during the M phase of the cell cycle, leading to the assumption that the majority of the transcriptome is then protected from active deadenylation [35].

2.4. Alternative Polyadenylation (APA)

Apart from the canonical PAS—the AAUAAA hexamer—other weaker signals called alternative PASs, may be present in transcripts. Generally, the higher the sequence similarity between an alternative and canonical PAS, the stronger the recognition of the alternative PAS. When this alternative PAS is selected as a signal for the cleavage and polyadenylation event, the process is described as alternative polyadenylation (APA). APA is thought to occur for around 70% of human protein-coding genes and can also affect non-coding RNAs, such as lncRNAs [37–39]. The affected transcripts can exhibit various numbers of APA events in a few or multiple APA sites. APA can dramatically modulate the expression of a specific gene and affect the fate of its transcript, including its half-life and cellular localization [40,41]. Depending on the alternative PAS localization, APA can occur either in 3'-UTRs (UTR-APA) or upstream of the last exon: in introns or protein-coding exons (UR-APA). UR-APA can lead to the production of truncated proteins with different functions (protein diversification), or the production of dysfunctional proteins. On the other hand, when APA occurs in the 3'-UTR of a transcript, it leads to the creation of an mRNA of different lengths, which still codes for a full-length protein. In such cases, APA can affect the expression of a gene by, for example, changing the number of miRNA-binding sites in the transcript. As it was shown that the 3'-UTR can regulate protein localization independently from mRNA localization, it can act as a scaffold for various protein complexes which, when recruited to translation sites, can interact with specific domains of newly translated proteins [42]. This, in turn, leads to the translocation of such proteins. An example is a CD47 transcript, whose short 3'-UTR promotes the localization of the protein at the ER, while its longer isoform promotes its translocation to the plasma membrane [37]. The occurrence of APA can be regulated in many ways, one being the 'strength' of alternative PASs. The more similar the sequence of an alternative PAS is to that of the canonical PAS, the stronger the alternative PAS will be. Moreover, the localization of a specific PAS within a transcript sequence is also worth mentioning. Typically, PASs localized closer to the start codon (proximal) are considered to be weaker, while PASs localized further from the start codon (distal) are stronger [5]. Core polyadenylation factors, as well as other RNA-binding proteins (RBPs), can also regulate APA. For example, PABPN1 enhances the selection of distal PASs by competing with cleavage and polyadenylation complexes: it recognizes a weak PAS and binds to it, thereby blocking CPSF binding [43,44]. Regarding RBPs, HuR protein favors the selection of a distal PAS by binding to U-rich elements lying close to a proximal PAS [45]. Another crucial group of RBPs involved in APA is muscleblind-like (MBNL) proteins, whose binding sites are present in the close vicinity of many PASs. In myotonic dystrophy (DM), MBNL proteins were shown to either activate or suppress polyadenylation at specific sites [46]. PABPN1, MBNL, and HuR proteins are described in more detail in the following chapters.

3. Methods for Determining the Length of Poly(A) Tails and Analyzing APA Sites

Here, we briefly summarize some of the available methods for measuring poly(A) tail lengths using both transcript-specific and transcriptome-wide approaches, as well as methods designed for APA site identification and the currently available APA databases.

3.1. Transcript-Specific Analyses of Poly(A) Tails

For transcript-specific analyses, RNase H/Oligo(dT) northern blotting was once the gold-standard technique. Oligo(dT) were hybridized with transcripts, followed by RNase H treatment. Next, poly(A) tail length was assessed through the comparison of transcripts with, or without, hybridized oligo(dT) using northern-blot-based detection [47]. One

advantage of this technique was that it avoided introducing a PCR-based bias. However, it had many disadvantages, one being the requirement of large amounts of RNA. This technique is also laborious and works best on highly expressed and shorter mRNAs. These disadvantages encouraged scientists to develop more effective and more precise methods that allow for the transcript-specific measurement of poly(A) tails. Such techniques are usually grouped together and described as poly(A) test (PAT) assays, which include PCR-based rapid amplification of cDNA ends PAT (RACE-PAT), ligation-mediated PAT (LM-PAT) [48], and extension PAT (ePAT) [49]. These techniques, as well as other assays, were reviewed extensively in [26,50]. Currently, the most popular assay is probably ePAT. Here, a 5'-anchored oligo(dT) primer serves as a template for Klenow polymerase, which extends the transcript on the basis of the anchor sequence. This oligo(dT) then serves as a reverse primer for cDNA synthesis. When cDNA is generated, this technique requires the use of a target-specific forward primer for PCR. The PCR products can then be visualized using agarose gel electrophoresis or capillary electrophoresis. Another variation of the ePAT method is the poly(A) tail-length assay, currently available as a commercial kit [51]. This assay uses PAP, which adds guanosine and inosine residues at the 3' end of every poly(A) tail that exists in the RNA sample. This added G/I tail then becomes a priming site for cDNA synthesis and the following PCR. These PAT methods are fast and easy to perform but require the optimization of the PCR. Otherwise, they may generate PCR-introduced biases and shorter products (i.e., those containing shorter poly(A) tails) may be amplified more efficiently.

3.2. Transcriptome-Wide Methods for Measurement of Poly(A) Tails

One method that can determine the lengths of poly(A) tails in both transcript-specific and transcriptome-wide manners is Nanopore technology sequencing [19,52]. Nanopore technology offers a set of products that allow for DNA and RNA sequencing, including for transcripts containing poly(A) tails. The principle of all Nanopore sequencing is that specific adaptors are attached to RNA or DNA, which then allow for such nucleic acids to be pulled through the pore at a constant rate. The length of a poly(A) tail is then estimated based on the correlation between length and the time spent being transferred through the pore. Thanks to numerous approaches, users can decide whether they want to use PCR amplification or bypass this step (to reduce PCR-introduced bias) and sequence the RNA or cDNA with only minimal library preparation, etc. This method is currently one of the best for measuring poly(A) tails and allows for the precise assignment of an analyzed poly(A) tail to a specific transcript. However, it requires large amounts of purified RNA and cannot identify non-A residues within the tail.

Other very popular methods for the transcriptome-wide measurement of poly(A) tail lengths include TAIL-Seq [53] and PAL-Seq [24]. Both methods are very similar in terms of their basic concepts: in the first step, RNAs lacking poly(A) tails (mainly rRNAs) are removed from a pool of sequenced transcripts. Moreover, biotinylated 3' adaptors are used to allow for the pull-down of bound transcripts on streptavidin beads, which is followed by the addition of 5' adaptors in both techniques. The main difference between TAIL-Seq and PAL-Seq is their sequencing methodology, which leads to slightly different results. TAIL-Seq allows for the determination of any non-A residues that might be present in the poly(A) tail (i.e., possible guanylation or uridylation). However, this technique is expensive and technically challenging. On the other hand, PAL-Seq permits the capture of tails with only adenosines present at their 3' ends, but requires less starting RNA material and, due to the enrichment of polyadenylated RNAs at the start of the procedure, is more efficient than the TAIL-Seq approach [19]. PAL-Seq also requires a highly specialized experimental setup and can only be performed on the Illumina platform. In light of these disadvantages, another method that combines the enrichment of polyadenylated RNA from PAL-Seq with the sequencing methodology and algorithm of the TAIL-Seq protocol was developed. This method was named mTAIL-Seq [54] and requires even less starting RNA than the two previously discussed assays. This method also provides higher sensitivity and is less costly than the previous methods; however, it has a limited ability to capture tails shorter

than 8 nt and tails that end with non-A nucleotides. Due to its drawbacks, mTAIL-Seq is predominantly used for measuring poly(A) tail lengths.

Recently, a new technique has been developed: full-length poly(A) and mRNA sequencing (FLAM-Seq) [55]. FLAM-Seq relies on 3' end RNA extension through G/I tailing, template switching, cDNA amplification, and PCR amplification. The transcript sequences are then determined through PacBio sequencing. It is a fast and simple method that allows for the generation of long reads and provides information on poly(A) tail length, sequences of multiple transcripts, and mRNA isoforms. An outstanding advantage of FLAM-Seq is that it allows for the sequencing of poly(A) tails with the detection of non-A residues inside the tail sequences. Moreover, it generates full-length sequences for thousands of mRNAs with a very low error rate [55].

3.3. APA-Identification Techniques and Databases

In discussing APA-identification techniques, we can distinguish between experimental methods and the various computational tools available (Table 1). An example of the former is a massive analysis of cDNA ends (MACE-Seq) [56]. This is an improved version of 3' single-end mRNA-Seq, based on the sequencing of a single molecule of transcripts. Other experimental methods facilitating APA detection include TAIL-Seq, PAL-Seq, and FLAM-Seq, all of which were described in the previous sub-chapter. Regarding computational APA-identification methods, APATrap allows for the identification and quantification of APA sites derived from various RNA-Seq data [57]. APATrap also allows for the identification of 3'-UTRs and extended 3'-UTRs, as well as for distinguishing genes that use various APA sites under different conditions. Other computational tools for detecting APA, such as DaPars [58], scDAPA [59], etc., were previously reviewed in detail by others [11,60].

Table 1. Summary of experimental methods and computational tools allowing for APA sites detection together with APA databases (accessed on 12 February 2022).

| Name | Year of Publication | Website | References |
|--|---------------------|---|------------|
| Computational tools for APA detection | | | |
| DaPars | 2014 | https://github.com/ZhengXia/dapars | [58] |
| Change point | 2014 | http://utr.sourceforge.net | [63] |
| Roar | 2016 | https://github.com/vodkatad/roar | [64] |
| APATrap | 2018 | https://apatrap.sourceforge.io | [57] |
| QAPA | 2018 | https://www.github.com/morrislab/qapa | [65] |
| TAPAS | 2018 | https://github.com/arefeen/TAPAS | [66] |
| KAPAC | 2018 | https://github.com/zavolanlab/PAQR_KAPAC.git | [67] |
| scDAPA | 2019 | https://scdapa.sourceforge.io | [59] |
| APA databases | | | |
| PolyA-Seq Atlas | 2012 | http://genome.ucsc.edu/ | [10] |
| APADB | 2014 | http://tools.genxpro.net:9000/apadb/ | [62] |
| APASdb | 2015 | http://genome.bucm.edu.cn/utr/ | [68] |
| PolyA_DB3 | 2018 | https://exon.apps.wistar.org/PolyA_DB/v3/ | [69] |
| APAAtlas | 2020 | https://hanlab.uth.edu/apa/ | [70] |
| PolyASite 2.0 | 2020 | https://polyasite.unibas.ch | [61] |
| Animal-APAdb | 2021 | http://gong_lab.hzau.edu.cn/Animal-APAdb/ | [71] |

Over recent years, several databases that gather information about possible APA sites in various transcripts have been created. The data collected there were obtained using the previously mentioned RNA-Seq methods and computational tools. One such database is PolyASite 2.0 [61]. It contains information about PASs based on data generated using numerous sequencing methods for human, mouse, and *C. elegans* transcriptomes [61]. Moreover, PolyASite 2.0 provides information about PASs located in 3'-UTRs, introns, and CDSs. Another database is APADB [62], which contains information about human, mouse, and chicken transcripts of both coding and non-coding genes. The data collected there were generated through the NGS-coupled 3' end sequencing of thousands of samples (e.g., MACE-Seq). APADB also provides information regarding the loss of bioinformatically predicted miRNA-binding sites, which may occur through APA events themselves. Other databases were well-reviewed recently [11,60].

4. Repeat Expansion Diseases

Repeat expansion diseases are a group of genetic disorders caused by expansions of polymorphic tracts of nucleotide sequences in specific human genes [72] (Table S1). There are more than 40 diseases that are classified as repeat expansion disorders, and this number continues to grow [72,73]. Most of them mainly affect the nervous system, but there are also some diseases that impact the muscular system. The most common expanded sequence motifs are trinucleotides, but longer ones, up to dodecanucleotide units, exist [74]. Repeat expansion diseases display several specific features: they are mostly autosomal dominant diseases, but recent reports indicate that mutations of expanded repeats also cause autosomal recessive diseases, e.g., cerebellar ataxia, neuropathy, and vestibular areflexia syndrome (CANVAS) [73]. Typically, patients are heterozygotes possessing wild-type and mutant alleles, with the latter harboring the expanded repeat tract. With regard to the mutant allele, usually the longer the repeat tract, the earlier the onset of symptoms and the more severe the disease course. However, the number of nucleotide repeats needed for disease onset differs among diseases. Expanded repeat tracts can occur in protein-coding sequences of particular genes, as well as in promoter sequences, introns, or in 5'- and 3'-UTRs (Table S1). Interestingly, many expanded repeat tracts in RNAs are prone to the formation of stable structures like hairpins or G-quadruplexes [75,76]. These structures affect transcript functioning, e.g., through sequestration of proteins [72], whereas many processes remain to be investigated in this context, including APA.

In this chapter, we shortly summarize the tissue-specific processing of poly(A) tails, focusing on the cell types most affected in repeat expansion diseases. Moreover, we discuss selected diseases in the context of APA and poly(A) tail lengths in the implicated mutant transcripts.

4.1. APA and Poly(A) Tail Lengths in Neurons and Muscles

An interesting aspect concerning polyadenylation in neurons is the occurrence of neuron-specific proteins that demonstrate a specific impact on polyadenylation. These proteins include the previously mentioned GLD2 [77], mammalian suppressor of tauopathy 2 (MSUT2) [78], zinc finger CCCH domain-containing protein 14 (ZC3H14) [79], and ataxin-2 (ATXN2) [80]. In terms of the present review, the most interesting protein is ATXN2, as the expanded repeat tract in ATXN2 is responsible for spinocerebellar ataxia type 2 (SCA2) [81], as well as normal ATXN2 is implicated in amyotrophic lateral sclerosis (ALS) [82]. Regarding polyadenylation, ATXN2 inhibits poly(A) nuclease (PAN) activity and causes the occurrence of longer poly(A) tails [80]. It was also suggested that ATXN2 might stabilize its associated mRNAs by suppressing their decay. Additionally, it may contribute to the APA of its associating mRNAs, thereby supporting activity-dependent translation and neural plasticity [80].

Over recent years, it has been shown that, for many genes, the selection of a particular polyadenylation site in a transcript is tissue-specific [83–85]. Regarding neurons, studies performed in *Drosophila*, mouse, and human cells demonstrated a neuron-specific enrichment of long 3'-UTRs in various transcripts [41,86–88]. An example of such a phenomenon is *Drosophila*'s mei-P26 transcript, whose 3'-UTR is very short in the testis and 18.5 kb long in the central nervous system (CNS) [86]. A tendency for longer 3'-UTR expression in the CNS was suggested to be limited to neurons and is not the case in astrocytes, microglia, or oligodendrocytes [58,63]. An example of a factor that favors the choice of distal PASs is the neuron-specific RBP Elav (embryonic lethal abnormal visual protein), first identified in *Drosophila* [88,89]. Elav binds to its target mRNA at proximal PASs, blocking cleavage and polyadenylation machinery, thereby promoting the use of distal PASs. In *Drosophila* embryos, in the absence of the Elav protein, several genes lack long 3'-UTRs [90]. Human homologs of the *Drosophila* Elav protein are the Hu-family proteins: HuR, HuB, HuC, and HuD [91].

Due to neuron-specific features such as their structure (a single axon can extend hundreds of centimeters in length) and function (the synaptic plasticity underlying learning and memory processes), a common feature of these cells is local translation [17,30,92]. This phenomenon is required at synapses to, for example, convert short-term memory

into long-term memory. Therefore, neurons have had to develop specific mechanisms for the transport and repression of the translation of transcripts (during their transport) throughout the cell, to its distal compartments (e.g., dendrites) [17,92]. Here, the α -subunit of calcium/calmodulin-dependent protein kinase II (α CaMKII) transcript serves as an example (reviewed in [17]). When needed in dendrites, this transcript is polyadenylated in the nucleus and exported to the soma, where it is deadenylated and repressed. It is transported in this state to dendrites and, upon neuronal stimulation, is polyadenylated again and locally translated. This is also an example of the regulation of poly(A) tail length in neurons, which depends on the cell's needs.

Regarding the muscular system, RNA-sequencing studies performed in *C. elegans* demonstrated the general use of APA sites in muscles. This research revealed that muscle tissue demonstrates a specific polyadenylation pattern [93]. An example of such a tissue-specific form of post-transcriptional regulation led by APA is the *PAX3* (Paired box gene-3) gene, which is a major regulator of muscle stem-cell development and is required for the formation of limb muscles [94] and in response to stress [95,96]. During myogenic differentiation, *PAX3* expression is silenced by miR-206 [97]. The APA-based use of proximal PAS in myogenic progenitors leads to 3'-UTR shortening of the *PAX3* transcript, which results in such mRNA escaping miRNA-induced gene silencing [97–99]. In muscles, RBPs also play an important role in the control of PAS selection and have an impact on the general polyadenylation profile. Among such RBPs we specify MBNL proteins, which are discussed in detail in the DM 1 subsection below.

4.2. Characteristics and Perturbations in Poly(A) Tail Processing in Repeat Expansion Diseases

4.2.1. Oculopharyngeal Muscular Dystrophy

A repeat expansion disease with crucial APA implications is dominant oculopharyngeal muscular dystrophy (OPMD), which is characterized by progressive eyelid drooping (ptosis), difficulty in swallowing, filamentous intranuclear inclusions in muscle fibers, and proximal limb weakness [100,101]. The gene encoding PABPN1, mentioned in the first chapter, is directly involved in this disease. The relevant mutation involves the expansion of a GCG repeat tract encoding the polyalanine tract, located at the N-terminus of the PABPN1. This protein plays an important role in RNA processing, e.g., it is required for efficient mRNA export from the nucleus [102] and controls the length of poly(A) tails [13]. The mutated version of PABPN1 was shown to sequester the WT protein in nuclear inclusions in the muscle fibers of OPMD patients, altering the function of the WT protein [103]. Moreover, it was recently suggested that the PABPN1 may act as an APA suppressor. This hypothesis was supported by the finding that this protein binds near proximal PASs and therefore suppresses the use of these PASs for cleavage and polyadenylation. Additionally, it was demonstrated that either the downregulation or mutation of PABPN1 resulted in the genome-wide use of proximal PASs, which finally cause 3'-UTR shortening to occur in both mouse models of OPMD and in human cells [43,44].

4.2.2. Fragile X-Associated Tremor/Ataxia Syndrome

Fragile X syndrome (FXS), and its related diseases-fragile X-associated immature ovarian insufficiency (FXPOI) and fragile X-associated tremor/ataxia syndrome (FXTAS)-are caused by the expansion of a CGG repeat tract located in the 5'-UTR of *fragile X mental retardation 1* gene (*FMR1*) [104]. In FXPOI and FXTAS, premutation alleles contain 55–200 CGG repeats; however, this mutation does not inactivate the gene [104]. On the other hand, in FXS, the full mutation in *FMR1* of more than 200 CGG repeats leads to transcriptional silencing of the gene [104]. Normally, *FMR1* encodes a protein (FMRP) which is important in brain development as it plays a role in the functioning of neuronal networks and neuronal plasticity [105,106]. In the 3'-UTR of an *FMR1* transcript, three PASs were identified: one canonical and two alternatives [107]. In premutation alleles, the level of transcript isoforms emerging from alternative PASs is decreased, indicating that *FMR1* premutation affects polyadenylation through APA phenomena and thus might

impact pathology. Additionally, it was shown that, in this case, APA produces mRNAs with different poly(A) tail lengths: a transcript with expanded CGG tract had a shorter poly(A) tail than its WT counterpart [107].

4.2.3. Amyotrophic Lateral Sclerosis

ALS is a neurodegenerative disease that causes the death of motor neurons that control skeletal muscle contraction. ALS progresses over time, leading to muscle stiffness, weakness, and finally, paralysis and death. This process takes 2–5 years from the initial appearance of symptoms. ALS can have a sporadic or familial character. To date, the pathogenesis of sporadic cases remains unclear, while familial ones are mainly caused by mutations in *C9orf72*, *SOD1*, *FUS*, or *TARDBP* [108,109].

The most common familial form of ALS is caused by the expansion of GGGGCC repeats in the first intron of the *C9orf72* gene. This intron should be degraded after splicing, but because of the mutation, it accumulates in the nucleus and forms RNA foci [110]. Alternatively, it is transported to the cytoplasm where it undergoes RAN translation which, in turn, creates potentially toxic polypeptides [108,111]. RNA-seq was used to determine the APA status of samples from ALS patients, and the results indicated that many defects exist in the overall polyadenylation of transcripts in ALS. Depending on the brain area and type of ALS, distal or proximal PASs are used: (I) in the case of the cerebellum in familial ALS, proximal, rather than distal, PASs are used, while in sporadic ALS, the use of proximal and distal PASs is almost equivalent; (II) in the frontal cortex in familial ALS, proximal and distal PASs have similar rates of use, while in sporadic ALS, distal PASs are used more than in proximal ones [112]. Similar studies demonstrated that 2.7% of genes have statistically significant deregulation of APA in ALS, with a tendency toward the extension of 3'-UTRs. Gene ontology analysis revealed that most of these genes are responsible for neuron-projection development and cytoskeletal intracellular transport [109]. In ALS, the deregulation of polyadenylation affects many genes and potentially seems to be part of the pathogenesis mechanism.

Other genes whose mutations cause ALS, *FUS*, and *TARDBP*, encode two RBPs: FUS and TDP-43 respectively. In *TARDBP*, the mutations can be localized in different regions, but the majority are in disordered low-complexity regions [113,114]. Moreover, in *FUS*, other mutations that cause ALS have been identified, with the most common being a substitution of R521 [114]. The localization of FUS and TDP-43 mutations in their nuclear localization signals (NLS) causes the depletion of these proteins from the nucleus and leads to their accumulation in inclusions in the cytoplasm. Both proteins control the alternative splicing of genes participating in neuronal development which, in the case of mutation, leads to neurodegeneration [115]. Moreover, FUS and TDP-43 are involved in the polyadenylation process. FUS can control the selection of the polyadenylation site depending on the distance from the FUS-binding site [115]. TDP-43's role is to repress cryptic exons, thereby stabilizing transcripts. TDP-43 suppresses a cryptic PAS of the *stathmin-2* mRNA, which encodes proteins crucial for the regeneration and growth of axons. The re-localization or depletion of TDP-43 causes the recognition of a premature PAS in the *stathmin-2* mRNA and leads to the production of non-functional transcripts [114]. TDP-43 also controls the PAS selection of its own transcript and, in this way, self-regulates its expression. In the case of an excess of TDP-43 in the nucleus, it inhibits the selection of the most proximal PAS (there are three alternative PASs within the *TARDBP* sequence), leading to the elongation of a transcript. In such cases, the splicing of additional introns occurs, which is followed by the export of the transcript to the cytoplasm and its degradation by nonsense-mediated mRNA decay. As a result, a decrease in the cytoplasmic level of *TARDBP* mRNA is observed. On the other hand, when there is less nuclear TDP-43, the proximal PAS is selected. In this case, the transcript is transported to the cytoplasm and translated. In ALS patients, mutant TDP-43 is depleted from the nucleus, which leads to an increase in the level of TDP-43 in motor neurons by the mechanism described above. This loop may accelerate the progress of the disease through the formation of cytoplasmic inclusions and disturbances in mRNA metabolism [116].

4.2.4. Myotonic Dystrophy Type 1

DM 1 is primarily characterized by muscle dystrophy or myotonia. A mutation responsible for DM1 is the expansion of CUG repeats present in the 3'-UTR of the *DMPK* gene. RNA containing the expanded CUG tract acts through gain of function mechanisms. Mutated transcripts are mainly localized in the nucleus, where they form RNA foci and sequester RBPs such as MBNL1, -2, and -3; HNRNP1; CUGBP1; and STAU1 [117,118]. This leads to the trans-deregulation of RNA metabolism—especially APA, miRNA processing, and alternative splicing processes [119,120]. MBNL's implications in alternative splicing and APA are well-documented [121–124]. The RNA-binding motif of this protein is composed of four zinc-finger domains [125]. Using this motif, MBNL binds to transcripts and act as an activator or repressor of splicing [126]. It was shown that MBNL1 can regulate alternative splicing in the brain, heart, and muscle in post-natal development [46,127]. In DM1, the sequestration of MBNL1 results in the functional inactivation of the protein, leading to disturbances in the alternative splicing of genes controlled by this protein [128,129]. Concerning APA, in MEF cells with *Mbnl1*, *Mbnl2*, and *Mbnl3* depletion, thousands of polyadenylation sites shifts were detected, which suggests that MBNL proteins act globally on alternative translation events [46]. The importance of MBNL1 sequestration in DM1 pathogenesis was demonstrated in skeletal muscle satellite cells [130]. These cells are responsible for muscle regeneration, and, in cases of DM1, this regeneration is impaired. Moreover, autophagy is increased, lowering the proliferation ability of skeletal satellite cells—a process that plays a key role in the initial steps of cell regeneration in damaged muscle. However, the genome modification of *DMPK*, or overexpression of the MBNL1 protein, can ameliorate proliferation defects in skeletal-muscle satellite cells. Indeed, it was revealed that increases in MBNL1 levels in the cytoplasm partially restored appropriate RNA processing [130].

4.2.5. Huntington's Disease

Huntington's disease (HD) is a neurodegenerative disorder caused by the expansion of the CAG repeat tract, encoding the polyglutamine (polyQ) tract, in the first exon of the huntingtin gene (*HTT*) [131,132]. The expansion is dominantly inherited, and patients usually develop symptoms in their mid-thirties. Patients harbor two alleles: the wild-type, containing 10 to 30 CAG repeats, and the mutant allele, with over 40 CAG repeats. The mutation of the CAG tract in *HTT* leads to the degeneration of neurons, which emerges in the motor cortex and striatum and is also substantial in other regions of the brain [132,133]. Huntingtin transcripts were reported to have several isoforms produced via APA. Three of these APA events occur in the 3'-UTR of the *HTT* gene, leading to three isoforms that differ only in their 3'-UTR length, with no impact on the coding sequence [134] (Figure 2a, three arrows marked in 3'-UTR). These 3'-UTR isoforms of the *HTT* transcript include a short isoform (PAS: AGUAAA), which is 10.3 kb long, and a long isoform (PAS: AUUAAA), which is 13.7 kb long [134,135]. The third one, 12.5 kb long, is described as an intermediate 3'-UTR isoform (PAS: AAUGAA) and, as well as the previous isoforms, is conserved between mice and humans [134]. This research also suggested that all of these *HTT* isoforms may have different half-lives and localization, as well as RBP- and miRNA-binding sites. Short and long 3'-UTR isoforms are considered to be the most common *HTT* isoforms: the short one is more abundant in actively dividing cells (e.g., B cells and the HEK 293 cell line) and peripheral tissues (muscles), while the long one predominates in non-dividing cells (i.e., terminally differentiated cells), such as in brain tissue, but also prevails in breast and ovary tissues [134,135]. Research using SH-SY5Y cells and focusing on poly(A) tail length in *HTT* transcripts revealed that each of these three *HTT* mRNA isoforms had different lengths of poly(A) tails. The short 3'-UTR isoform possesses around 60 adenosine residues in its poly(A) tail, while the intermediate and long 3'-UTR isoforms have around 5 and 10 A residues, respectively [134]. These results suggested that *HTT* transcripts have rather short poly(A) tail lengths, at least in this specific cell line. It was also shown that these isoforms arise from both alleles (normal and mutant) in HD, meaning that CAG expansion

probably does not have an impact on PAS selection in the 3'-UTR of *HTT* [134]. However, another study concluded that the short 3'-UTR isoform might be translated more efficiently than the longer one [136].

In HD, an additional isoform of the transcript (*HTT**Exon1*) is produced predominantly from mutant *HTT* due to APA site selection in intron 1 [137] (Figure 2a, red arrow in the first intron of *HTT*). This APA event results in an isoform of 7.9 kb, which is translated into the highly truncated huntingtin protein, with a high tendency to form aggregates. This mRNA isoform has been detected in HD patients but not in healthy individuals [137]. Another study revealed higher levels of this isoform in the HD patient hippocampus and motor cortex compared with controls [138]. Those results strongly suggest that the production of *HTT* mRNA containing exon 1 and part of intron 1 is associated with HD. One hypothesis that might explain this generation of such an mRNA isoform is related to the serine/arginine-rich splicing factor 6 (SRSF6). Bioinformatic analysis predicted that SRSF6 binds to CAG and CAGCAA repeats, resulting in the incomplete splicing of exon 1 to exon 2 (intron 1 retention) and simultaneously leading to the exposure of cryptic poly(A) sites within intron 1 [139]. However, a subsequent *in vivo* study performed by the same group demonstrated that SRSF6 is not needed for the aberrant splicing of *HTT* [140], meaning that this protein and its sequestration by CAG repeats is not responsible for a mechanism leading to the formation of *HTT**Exon1* mRNA, which warrants further investigation.

4.3. Potential Perturbations in Poly(A) Tail Processing in Other PolyQ Diseases

HD is one of nine polyQ diseases that are all caused by CAG repeat expansion located in the ORF regions of specific genes [141]. Due to already-known perturbations in poly(A) tail processing in *HTT*, as described in the previous subchapter, we more closely examined other transcripts directly involved in polyQ diseases. Apart from *HTT*, we searched the PolyASite 2.0 database [61] for information concerning alternative polyadenylation sites in eight other mRNAs: *AR*, *ATN1*, *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, and *TBP*, whose genes contain mutations responsible for spinal and bulbar muscular atrophy (SBMA), dentatorubral-pallidoluysian atrophy (DRPLA), and spinocerebellar ataxia types 1, 2, 3, 6, 7, and 17 (SCA1, SCA2, SCA3, SCA6, SCA7, and SCA17, respectively). For all the mRNAs, 3' end sequencing data were available and included a set of sites for alternative polyadenylation (Figure 2b–i). Apart from 3'-UTRs, identified PASs were located in protein-coding exons and introns, but the use of these sites was often quantified as very low. Nevertheless, the identification of these sites suggests the possibility of premature transcription termination (PTT) [142] for these transcripts which may specifically occur under certain conditions. Selected sites are presented in Figure 2, where we include a presentation of PASs with relatively high TPM values (≥ 0.1) within all the gene regions, as well as all the PASs that fulfilled at least one of the following criteria: (I) presence in the same exon as a CAG repeat tract, (II) presence in an exonic sequence at a distance of less than 500 nt from a CAG repeat tract, or (III) presence in introns directly adjacent to an exon containing CAG repeat tract. The purpose of these criteria was to verify the presence of PASs that could potentially be of importance in cases of repeat-tract mutations. Indeed, numerous PASs were identified according to the listed criteria for most analyzed mRNAs. Abnormal interactions of proteins with expanded CAG repeats in RNA, or the abnormal secondary structure of this mutated region [143], can potentially lead to disruptions in splicing and/or PAS recognition. In several of the investigated transcripts, PASs were present in neighboring introns, which could lead to the production of a shorter transcript and smaller protein if a splicing aberration—similar to that reported for *HTT*—occurs. These effects may be crucial for pathogenesis, especially for mRNAs where the CAG repeat tract is located closer to the 5' end of a transcript, such as in *HTT*. For mRNAs including *AR*, *ATXN2*, and *TBP*, if the resulting transcript would still contain the mutant CAG tract, it would be translated into a very short aberrant protein with expanded polyQ tracts that could be very prone to aggregation. These events are speculative at present, but, as was shown in HD, may be important contributors in cascades of pathogenic events.

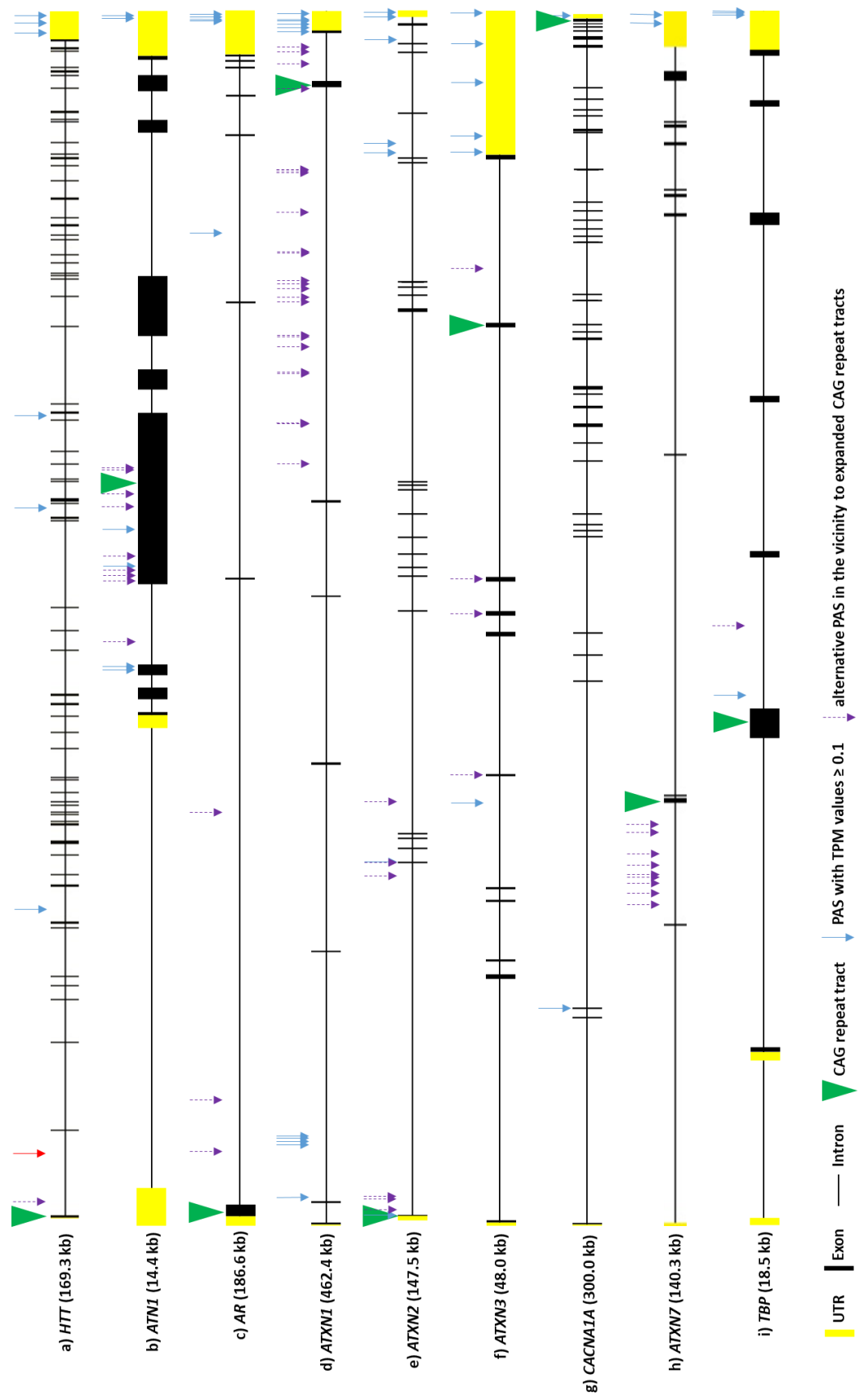


Figure 2. A scheme representing genes associated with polyQ diseases, based on Ensembl data [144], with marked alternative polyadenylation sites derived from PolyASite 2.0 database. The red arrow represents PAS identified in intron 1 of mutant *HTT* gene. See text for more details.

In a set of analyzed mRNAs, there exists variation in the number and distribution of PASs in the 3'-UTRs (Figure 2). Most of the transcripts have more than one PAS identified in the 3'-UTR. As previously mentioned, one of the consequences of the selection of proximal or distal PASs, especially in the 3'-UTR, is the presence or absence of miRNA-binding sites. We included additional analysis of the data gathered in APAdb [62] to list miRNAs that lose their binding sites in the case of proximal or intermediate PAS selection in *HTT* (Figure 3a) and *ATXN3* (Figure 3b) mRNAs. Numerous miRNAs are identified as potentially regulating these transcripts and potential proximal PAS selection will lead to a lack of these expression-regulating sites.

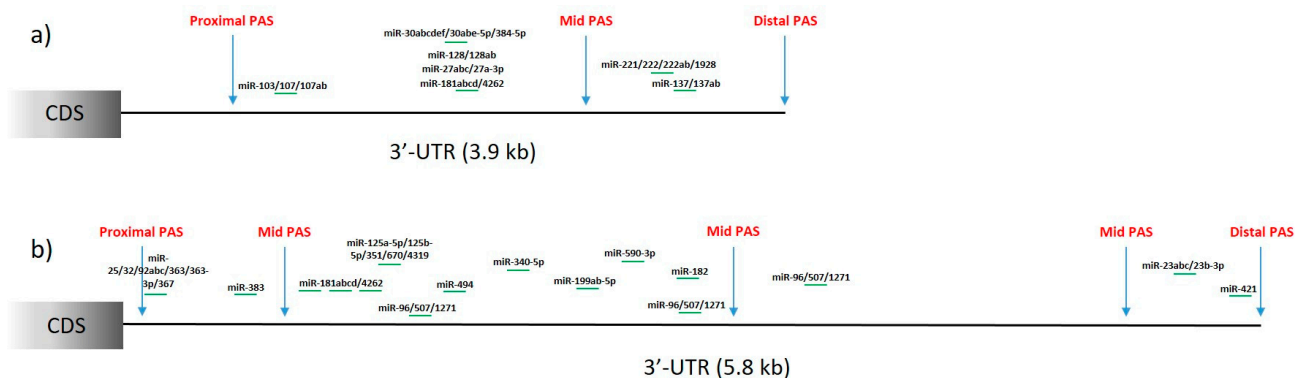


Figure 3. A schematic representation of miRNA-binding sites that are present in 3'-UTRs of *HTT* (a) and *ATXN3* (b) mRNAs with available PASs marked. Selection of PAS other than distal will lead to a lack of specific miRNA-binding sites.

5. Conclusions and Perspectives

3'-UTRs, where polyadenylation events usually occur, had their lengths generally increased during evolution and therefore are considered as one of the elements of increased molecular, and as a result also morphological, complexity [145]. APA adds another level for the regulatory role of 3'-UTRs, as transcripts that substantially differ in 3'-UTR length can be generated. This results in the presence, or absence, of specific structural elements, RNA- and protein-binding regions, etc., and thus it affects the metabolism of a given transcript [146].

APA has various biological functions, both physiological and pathological. The most recent data suggest that APA can be a molecular tool for regulating mRNA and protein levels during cellular responses, e.g., differentiation [38]. On the other hand, as APA can affect the expression of various genes and can produce incorrect proteins, it is also associated with many diseases, whether oncological, immunological, endocrine, or neurological.

5.1. Methodological Aspects

As poly(A) tail processing is involved in many physiological and pathological mechanisms, there is still a great need for the development of various tools and methods that will allow for more insightful studies of APA sites and tail lengths. It has now become possible to precisely analyze polyadenylation at the single-cell level [41,147,148], but data gathered from multiple cell types and at different developmental stages is needed to advance our understanding of events connected to polyadenylation. Moreover, this process is clearly associated with other molecular mechanisms that occur for specific transcripts, and a further consolidation of the available data is required to provide a comprehensive view of transcript functioning.

5.2. Implications for Pathogenesis of Repeat Expansion Diseases

As the majority of transcripts undergo the selection of APA sites, it remains to be elucidated how eventual mis-selection events are connected with diseases. In mutant

transcripts, repeat tract expansion can affect PAS selection, as was described in HD. Repeat expansion diseases mainly affect specific cell types, e.g., a specific vulnerability of neuronal cells to degeneration is observed [149]. Gathering more and more data from disease tissues and cells will enable the identification of potential disruptions which could occur in specific cell types in cases of repeat tract mutation. Additionally, as widespread APA deregulation has been already observed in a range of neurological disorders [109], it would be of advantage to establish what are the specifically deregulated APA events also in repeat expansion diseases. This knowledge could be used for molecular diagnostics, as it has been recently suggested that APA-signatures in cancer appear to outperform existing biomarkers [58,150,151].

5.3. Implications into Therapy of Repeat Expansion Diseases

All repeat expansion diseases are as yet incurable, and often, especially for diseases where the expanded repeat tract is located in a coding sequence, the most straightforward therapeutic approach would be to eliminate the expression of the mutant gene. Many challenges remain to be faced in the development of these strategies. A crucial aim is to advance our knowledge regarding the therapeutic target features that could allow for the design of the most efficient and safest molecules for therapy. Mutant transcripts are currently the most promising targets, as various tools can be designed to bind to them in order to inhibit mutant gene expression. In regards to the therapeutic aspects, it remains to be determined whether specific 3' end transcript variants are present in cases of repeat tract mutation, and, if so, how they could be specifically targeted. Specific therapeutic interventions based on modulation of biogenesis or direct elimination of faulty APA isoforms are being currently extensively developed for other diseases such as cancer (as reviewed in [3]). Currently, the most desirable therapeutic tools for the majority of repeat expansion diseases are allele-selective ones, which preferentially target the expression of the mutant allele. Various types of oligonucleotides-including short interfering RNAs (siRNAs) and antisense oligonucleotides (ASOs)-have been tested in animal models, and clinical testing is ongoing [152]. Interestingly, some of the mutant mRNA-targeting oligonucleotides may cause deadenylation. Our recent study, focusing on the mechanisms leading to the allele-selective silencing of *HTT*, demonstrated that a specific self-duplexing siRNA (sd-siRNA) [153], which is described as miRNA-like siRNA, caused the rapid deadenylation of the mutant transcript and translational inhibition [154]. It remains to be investigated how deadenylation is activated in this therapeutic strategy using an endogene of *HTT*.

It seems we are greatly advancing our understanding of how important the poly(A) tail-connected events are in the molecular biology puzzles. In repeat expansion diseases, we already know some relationships of APA to selected diseases, but it looks like a lot is ahead of us to unravel.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cells11040677/s1>, Table S1: Repeat expansion diseases.

Author Contributions: General conceptualization by A.F. and A.C.; P.J. wrote the majority of manuscript draft and prepared figures and tables. G.F., A.F. and A.C. wrote some fragments and provided comments for manuscript draft. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Centre (2015/19/B/NZ2/02453).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

| | |
|----------------|------------------------------------|
| ALS | amyotrophic lateral sclerosis |
| APA | alternative polyadenylation |
| DM | myotonic dystrophy |
| HD | Huntington's disease |
| OPMD | oculopharyngeal muscular dystrophy |
| PABPN1 | poly(A) binding protein nuclear 1 |
| PAS | polyadenylation signal |
| PAT | poly(A) test |
| polyQ diseases | polyglutamine diseases |

References

- Casañal, A.; Kumar, A.; Hill, C.H.; Easter, A.D.; Emsley, P.; Degliesposti, G.; Gordiyenko, Y.; Santhanam, B.; Wolf, J.; Wiederhold, K.; et al. Architecture of eukaryotic mRNA 3'-end processing machinery. *Science* **2017**, *358*, 1056–1059. [[CrossRef](#)] [[PubMed](#)]
- Passmore, L.A.; Collier, J. Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. *Nat. Rev. Mol. Cell Biol.* **2021**. [[CrossRef](#)] [[PubMed](#)]
- Nourse, J.; Spada, S.; Danckwardt, S. Emerging roles of RNA 3'-end cleavage and polyadenylation in pathogenesis, diagnosis and therapy of human disorders. *Biomolecules* **2020**, *10*, 915. [[CrossRef](#)] [[PubMed](#)]
- Stewart, M. Polyadenylation and nuclear export of mRNAs. *J. Biol. Chem.* **2019**, *294*, 2977–2987. [[CrossRef](#)] [[PubMed](#)]
- Di Giammartino, D.C.; Nishida, K.; Manley, J.L. Mechanisms and Consequences of Alternative Polyadenylation. *Mol. Cell* **2011**, *43*, 853–866. [[CrossRef](#)]
- Sadek, J.; Omer, A.; Hall, D.; Ashour, K.; Gallouzi, I.E. Alternative polyadenylation and the stress response. *Wiley Interdiscip. Rev. RNA* **2019**, *10*, 1–16. [[CrossRef](#)]
- Ren, F.; Zhang, N.; Zhang, L.; Miller, E.; Pu, J.J. Alternative Polyadenylation: A new frontier in post transcriptional regulation. *Biomark. Res.* **2020**, *8*, 67. [[CrossRef](#)]
- Danckwardt, S.; Hentze, M.W.; Kulozik, A.E. 3' end mRNA processing: Molecular mechanisms and implications for health and disease. *EMBO J.* **2008**, *27*, 482–498. [[CrossRef](#)]
- Proudfoot, N.J. Ending the message: Poly(A) signals then and now. *Genes Dev.* **2011**, *25*, 1770–1782. [[CrossRef](#)]
- Derti, A.; Garrett-Engle, P.; MacIsaac, K.D.; Stevens, R.C.; Sriram, S.; Chen, R.; Rohl, C.A.; Johnson, J.M.; Babak, T. A quantitative atlas of polyadenylation in five mammals. *Genome Res.* **2012**, *22*, 1173–1183. [[CrossRef](#)]
- Gruber, A.J.; Zavolan, M. Alternative cleavage and polyadenylation in health and disease. *Nat. Rev. Genet.* **2019**, *20*, 599–614. [[CrossRef](#)] [[PubMed](#)]
- Zhao, J.; Hyman, L.; Moore, C. Formation of mRNA 3' Ends in Eukaryotes: Mechanism, Regulation, and Interrelationships with Other Steps in mRNA Synthesis. *Microbiol. Mol. Biol. Rev.* **1999**, *63*, 405–445. [[CrossRef](#)]
- Eckmann, C.R.; Rammelt, C.; Wahle, E. Control of poly(A) tail length. *Wiley Interdiscip. Rev. RNA* **2011**, *2*, 348–361. [[CrossRef](#)] [[PubMed](#)]
- Natalizio, B.J.; Wenthe, S.R. Postage for the messenger: Designating routes for nuclear mRNA export. *Trends Cell Biol.* **2013**, *23*, 365–373. [[CrossRef](#)] [[PubMed](#)]
- Bresson, S.M.; Conrad, N.K. The Human Nuclear Poly(A)-Binding Protein Promotes RNA Hyperadenylation and Decay. *PLoS Genet.* **2013**, *9*, 1003893. [[CrossRef](#)]
- Eisen, T.J.; Eichhorn, S.W.; Subtelny, A.O.; Lin, K.S.; McGeary, S.E.; Gupta, S.; Bartel, D.P. The Dynamics of Cytoplasmic mRNA Metabolism. *Mol. Cell* **2020**, *77*, 786–799. [[CrossRef](#)]
- Weill, L.; Belloc, E.; Bava, F.-A.; Méndez, R. Translational control by changes in poly(A) tail length: Recycling mRNAs. *Nat. Struct. Mol. Biol.* **2012**, *19*, 577–585. [[CrossRef](#)]
- Xiang, K.; Bartel, D.P. The molecular basis of coupling between poly(A)-tail length and translational efficiency. *eLife* **2021**, *10*, 66493. [[CrossRef](#)]
- Nicholson, A.L.; Pasquinelli, A.E. Tales of Detailed Poly(A) Tails. *Trends Cell Biol.* **2019**, *29*, 191–200. [[CrossRef](#)]
- Dominski, Z.; Yang, X.-C.; Kaygun, H.; Dadlez, M.; Marzluff, W.F. A 3' Exonuclease that Specifically Interacts with the 3' End of Histone mRNA. *Mol. Cell* **2003**, *12*, 295–305. [[CrossRef](#)]
- Lorenzi, L.; Chiu, H.S.; Avila Cobos, F.; Gross, S.; Volders, P.J.; Cannoodt, R.; Nuytens, J.; Vanderheyden, K.; Anckaert, J.; Lefever, S.; et al. The RNA Atlas expands the catalog of human non-coding RNAs. *Nat. Biotechnol.* **2021**, *39*, 1453–1465. [[CrossRef](#)]
- Chen, L.L. Linking Long Noncoding RNA Localization and Function. *Trends Biochem. Sci.* **2016**, *41*, 761–772. [[CrossRef](#)]
- Lima, S.A.; Chipman, L.B.; Nicholson, A.L.; Chen, Y.H.; Yee, B.A.; Yeo, G.W.; Collier, J.; Pasquinelli, A.E. Short poly(A) tails are a conserved feature of highly expressed genes. *Nat. Struct. Mol. Biol.* **2017**, *24*, 1057–1063. [[CrossRef](#)]
- Subtelny, A.O.; Eichhorn, S.W.; Chen, G.R.; Sive, H.; Bartel, D.P. Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature* **2014**, *508*, 66–71. [[CrossRef](#)]
- Gu, H.; Gupta, J.D.; Schoenberg, D.R. The poly(A)-limiting element is a conserved cis-acting sequence that regulates poly(A) tail length on nuclear pre-mRNAs. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 8943–8948. [[CrossRef](#)] [[PubMed](#)]

26. Jalkanen, A.L.; Coleman, S.J.; Wilusz, J. Determinants and implications of mRNA poly(A) tail size—Does this protein make my tail look big? *Semin. Cell Dev. Biol.* **2014**, *34*, 24–32. [[CrossRef](#)] [[PubMed](#)]
27. Charlesworth, A.; Meijer, H.A.; de Moor, C.H. Specificity factors in cytoplasmic polyadenylation. *Wiley Interdiscip. Rev. RNA* **2013**, *4*, 437. [[CrossRef](#)] [[PubMed](#)]
28. Ivshina, M.; Lasko, P.; Richter, J.D. Cytoplasmic polyadenylation element binding proteins in development, health, and disease. *Annu. Rev. Cell Dev. Biol.* **2014**, *30*, 393–415. [[CrossRef](#)]
29. Mendez, R.; Richter, J.D. Translational control by CPEB: A means to the end. *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 521–529. [[CrossRef](#)]
30. Udagawa, T.; Swanger, S.A.; Takeuchi, K.; Kim, J.H.; Nalavadi, V.; Shin, J.; Lorenz, L.J.; Zukin, R.S.; Bassell, G.J.; Richter, J.D. Bidirectional Control of mRNA Translation and Synaptic Plasticity by the Cytoplasmic Polyadenylation Complex. *Mol. Cell* **2012**, *47*, 253–266. [[CrossRef](#)]
31. Lim, J.; Kim, D.; Lee, Y.S.; Ha, M.; Lee, M.; Yeo, J.; Chang, H.; Song, J.; Ahn, K.; Kim, V.N. Mixed tailing by TENT4A and TENT4B shields mRNA from rapid deadenylation. *Science* **2018**, *361*, 701–704. [[CrossRef](#)] [[PubMed](#)]
32. Lim, J.; Ha, M.; Chang, H.; Kwon, S.C.; Simanshu, D.K.; Patel, D.J.; Kim, V.N. Uridylation by TUT4 and TUT7 marks mRNA for degradation. *Cell* **2014**, *159*, 1365–1376. [[CrossRef](#)]
33. Morgan, M.; Much, C.; DiGiacomo, M.; Azzi, C.; Ivanova, I.; Vitsios, D.M.; Pistolic, J.; Collier, P.; Moreira, P.N.; Benes, V.; et al. mRNA 3' uridylation and poly(A) tail length sculpt the mammalian maternal transcriptome. *Nature* **2017**, *548*, 347–351. [[CrossRef](#)] [[PubMed](#)]
34. Kojima, S.; Sher-Chen, E.L.; Green, C.B. Circadian control of mRNA polyadenylation dynamics regulates rhythmic protein expression. *Genes Dev.* **2012**, *26*, 2724–2736. [[CrossRef](#)] [[PubMed](#)]
35. Liu, Y.; Nie, H.; Lu, F. Dynamic RNA 3' Uridylation and Guanylation during Mitosis. *iScience* **2020**, *23*. [[CrossRef](#)]
36. Park, J.E.; Yi, H.; Kim, Y.; Chang, H.; Kim, V.N. Regulation of Poly(A) Tail and Translation during the Somatic Cell Cycle. *Mol. Cell* **2016**, *62*, 462–471. [[CrossRef](#)] [[PubMed](#)]
37. Tian, B.; Manley, J.L. Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol. Cell Biol.* **2016**, *18*, 18–30. [[CrossRef](#)]
38. Sommerkamp, P.; Cabezas-Wallscheid, N.; Trumpp, A. Alternative Polyadenylation in Stem Cell Self-Renewal and Differentiation. *Trends Mol. Med.* **2021**, *27*, 660–672. [[CrossRef](#)]
39. Elkon, R.; Ugalde, A.P.; Agami, R. Alternative cleavage and polyadenylation: Extent, regulation and function. *Nat. Rev. Genet.* **2013**, *14*, 496–506. [[CrossRef](#)]
40. Lutz, C.S.; Moreira, A. Alternative mRNA polyadenylation in eukaryotes: An effective regulator of gene expression. *Wiley Interdiscip. Rev. RNA* **2011**, *2*, 22–31. [[CrossRef](#)]
41. Agarwal, V.; Lopez-Darwin, S.; Kelley, D.R.; Shendure, J. The landscape of alternative polyadenylation in single cells of the developing mouse embryo. *Nat. Commun.* **2021**, *12*, 5101. [[CrossRef](#)]
42. Berkovits, B.D.; Mayr, C. Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. *Nature* **2015**, *522*, 363–367. [[CrossRef](#)] [[PubMed](#)]
43. Jenal, M.; Elkon, R.; Loayza-Puch, F.; Van Haaften, G.; Kühn, U.; Menzies, F.M.; Vrieling, J.A.F.O.; Bos, A.J.; Drost, J.; Rooijers, K.; et al. The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* **2012**, *149*, 538–553. [[CrossRef](#)]
44. Batra, R.; Manchanda, M.; Swanson, M.S. Global insights into alternative polyadenylation regulation. *RNA Biol.* **2015**, *12*, 597–602. [[CrossRef](#)]
45. Dai, W.; Zhang, G.; Makeyev, E.V. RNA-binding protein HuR autoregulates its expression by promoting alternative polyadenylation site usage. *Nucleic Acids Res.* **2012**, *40*, 787–800. [[CrossRef](#)] [[PubMed](#)]
46. Batra, R.; Charizanis, K.; Manchanda, M.; Mohan, A.; Li, M.; Finn, D.J.; Goodwin, M.; Zhang, C.; Sobczak, K.; Thornton, C.A.; et al. Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol. Cell* **2014**, *56*, 311–322. [[CrossRef](#)] [[PubMed](#)]
47. Sippel, A.E.; Stavrianopoulos, J.G.; Schutz, G.; Feigelson, P. Translational properties of rabbit globin mRNA after specific removal of poly(A) with ribonuclease H. *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 4635–4639. [[CrossRef](#)]
48. Sallés, F.J.; Richards, W.G.; Strickland, S. Assaying the polyadenylation state of mRNAs. *Methods* **1999**, *17*, 38–45. [[CrossRef](#)]
49. Jänicke, A.; Vancuylenberg, J.; Boag, P.R.; Traven, A.; Beilharz, T.H. ePAT: A simple method to tag adenylated RNA to measure poly(A)-tail length and other 3' RACE applications. *RNA* **2012**, *18*, 1289–1295. [[CrossRef](#)]
50. Murray, E.L.; Schoenberg, D.R. Chapter 24 Assays for Determining Poly(A) Tail Length and the Polarity of mRNA Decay in Mammalian Cells. *Methods Enzymol.* **2008**, *448*, 483–504.
51. Kusov, Y.Y.; Shatirishvili, G.; Dzagurov, G.; Gauss-Müller, V. A new G-tailing method for the determination of the poly(A) tail length applied to hepatitis A virus RNA. *Nucleic Acids Res.* **2001**, *29*, e57. [[CrossRef](#)] [[PubMed](#)]
52. Garalde, D.R.; Snell, E.A.; Jachimowicz, D.; Sipos, B.; Lloyd, J.H.; Bruce, M.; Pantic, N.; Admassu, T.; James, P.; Warland, A.; et al. Highly parallel direct RNA sequencing on an array of nanopores. *Nat. Methods* **2018**, *15*, 201–206. [[CrossRef](#)]
53. Chang, H.; Lim, J.; Ha, M.; Kim, V.N. TAIL-seq: Genome-wide determination of poly(A) tail length and 3' end modifications. *Mol. Cell* **2014**, *53*, 1044–1052. [[CrossRef](#)] [[PubMed](#)]
54. Lim, J.; Lee, M.; Son, A.; Chang, H.; Kim, V.N. MTAIL-seq reveals dynamic poly(A) tail regulation in oocyte-to-embryo development. *Genes Dev.* **2016**, *30*, 1671–1682. [[CrossRef](#)] [[PubMed](#)]

55. Legnini, I.; Alles, J.; Karaiskos, N.; Ayoub, S.; Rajewsky, N. FLAM-seq: Full-length mRNA sequencing reveals principles of poly(A) tail length control. *Nat. Methods* **2019**, *16*, 879–886. [[CrossRef](#)] [[PubMed](#)]
56. Tushev, G.; Glock, C.; Heumüller, M.; Biever, A.; Jovanovic, M.; Schuman, E.M. Alternative 3' UTRs Modify the Localization, Regulatory Potential, Stability, and Plasticity of mRNAs in Neuronal Compartments. *Neuron* **2018**, *98*, 495–511.e6. [[CrossRef](#)] [[PubMed](#)]
57. Ye, C.; Long, Y.; Ji, G.; Li, Q.Q.; Wu, X. APAttrap: Identification and quantification of alternative polyadenylation sites from RNA-seq data. *Bioinformatics* **2018**, *34*, 1841–1849. [[CrossRef](#)]
58. Xia, Z.; Donehower, L.A.; Cooper, T.A.; Neilson, J.R.; Wheeler, D.A.; Wagner, E.J.; Li, W. Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types. *Nat. Commun.* **2014**, *5*, 5274. [[CrossRef](#)]
59. Ye, C.; Zhou, Q.; Wu, X.; Yu, C.; Ji, G.; Saban, D.R.; Li, Q.Q. ScDAPA: Detection and visualization of dynamic alternative polyadenylation from single cell RNA-seq data. *Bioinformatics* **2020**, *36*, 1262–1264. [[CrossRef](#)]
60. Zhang, Y.; Liu, L.; Qiu, Q.; Zhou, Q.; Ding, J.; Lu, Y.; Liu, P. Alternative polyadenylation: Methods, mechanism, function, and role in cancer. *J. Exp. Clin. Cancer Res.* **2021**, *40*, 51. [[CrossRef](#)]
61. Herrmann, C.J.; Schmidt, R.; Kanitz, A.; Artimo, P.; Gruber, A.J.; Zavolan, M. PolyASite 2.0: A consolidated atlas of polyadenylation sites from 3' end sequencing. *Nucleic Acids Res.* **2020**, *48*, D174–D179. [[CrossRef](#)] [[PubMed](#)]
62. Müller, S.; Rycak, L.; Afonso-Grunz, F.; Winter, P.; Zawada, A.M.; Damrath, E.; Scheider, J.; Schmäh, J.; Koch, I.; Kahl, G.; et al. APADB: A database for alternative polyadenylation and microRNA regulation events. *Database* **2014**, *2014*, 1–11. [[CrossRef](#)] [[PubMed](#)]
63. Wang, W.; Wei, Z.; Li, H. A change-point model for identifying 3'UTR switching by next-generation RNA sequencing. *Bioinformatics* **2014**, *30*, 2162–2170. [[CrossRef](#)] [[PubMed](#)]
64. Grassi, E.; Mariella, E.; Lembo, A.; Molineris, I.; Provero, P. Roar: Detecting alternative polyadenylation with standard mRNA sequencing libraries. *BMC Bioinform.* **2016**, *17*, 423. [[CrossRef](#)] [[PubMed](#)]
65. Ha, K.C.H.; Blencowe, B.J.; Morris, Q. QAPA: A new method for the systematic analysis of alternative polyadenylation from RNA-seq data. *Genome Biol.* **2018**, *19*, 45. [[CrossRef](#)]
66. Arefeen, A.; Liu, J.; Xiao, X.; Jiang, T. TAPAS: Tool for alternative polyadenylation site analysis. *Bioinformatics* **2018**, *34*, 2521–2529. [[CrossRef](#)]
67. Gruber, A.J.; Schmidt, R.; Ghosh, S.; Martin, G.; Gruber, A.R.; van Nimwegen, E.; Zavolan, M. Discovery of physiological and cancer-related regulators of 3' UTR processing with KAPAC. *Genome Biol.* **2018**, *19*, 44. [[CrossRef](#)]
68. You, L.; Wu, J.; Feng, Y.; Fu, Y.; Guo, Y.; Long, L.; Zhang, H.; Luan, Y.; Tian, P.; Chen, L.; et al. APASdb: A database describing alternative poly(A) sites and selection of heterogeneous cleavage sites downstream of poly(A) signals. *Nucleic Acids Res.* **2015**, *43*, D59–D67. [[CrossRef](#)]
69. Wang, R.; Nambiar, R.; Zheng, D.; Tian, B. PolyA-DB 3 catalogs cleavage and polyadenylation sites identified by deep sequencing in multiple genomes. *Nucleic Acids Res.* **2018**, *46*, D315–D319. [[CrossRef](#)]
70. Hong, W.; Ruan, H.; Zhang, Z.; Ye, Y.; Liu, Y.; Li, S.; Jing, Y.; Zhang, H.; Diao, L.; Liang, H.; et al. APAAtlas: Decoding alternative polyadenylation across human tissues. *Nucleic Acids Res.* **2020**, *48*, D34–D39. [[CrossRef](#)]
71. Jin, W.; Zhu, Q.; Yang, Y.; Yang, W.; Wang, D.; Yang, J.; Niu, X.; Yu, D.; Gong, J. Animal-APAdb: A comprehensive animal alternative polyadenylation database. *Nucleic Acids Res.* **2020**, *49*, 47–54. [[CrossRef](#)] [[PubMed](#)]
72. Malik, I.; Kelley, C.P.; Wang, E.T.; Todd, P.K. Molecular mechanisms underlying nucleotide repeat expansion disorders. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 589–607. [[CrossRef](#)] [[PubMed](#)]
73. Ishiura, H.; Tsuji, S. Advances in repeat expansion diseases and a new concept of repeat motif–phenotype correlation. *Curr. Opin. Genet. Dev.* **2020**, *65*, 176–185. [[CrossRef](#)]
74. Paulson, H. Repeat Expansion Diseases. *Handb. Clin. Neurol.* **2018**, *147*, 105–123. [[PubMed](#)]
75. Krzyzosiak, W.J.; Sobczak, K.; Wojciechowska, M.; Fiszler, A.; Mykowska, A.; Kozłowski, P. Triplet repeat RNA structure and its role as pathogenic agent and therapeutic target. *Nucleic Acids Res.* **2012**, *40*, 11–26. [[CrossRef](#)] [[PubMed](#)]
76. Pan, F.; Zhang, Y.; Xu, P.; Man, V.H.; Roland, C.; Wenginger, K.; Sagui, C. Molecular conformations and dynamics of nucleotide repeats associated with neurodegenerative diseases: Double helices and CAG hairpin loops. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 2819–2832. [[CrossRef](#)]
77. Jae, E.K.; Drier, E.; Barbee, S.A.; Ramaswami, M.; Yin, J.C.P.; Wickens, M. GLD2 poly(A) polymerase is required for long-term memory. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 14644–14649. [[CrossRef](#)]
78. Wheeler, J.M.; McMillan, P.; Strovast, T.J.; Liachko, N.F.; Amlie-Wolf, A.; Kow, R.L.; Klein, R.L.; Szot, P.; Robinson, L.; Guthrie, C.; et al. Activity of the poly(A) binding protein MSUT2 determines susceptibility to pathological tau in the mammalian brain. *Sci. Transl. Med.* **2019**, *11*, aao6545. [[CrossRef](#)]
79. Morris, K.J.; Corbett, A.H. The polyadenosine RNA-binding protein ZC3H14 interacts with the THO complex and coordinately regulates the processing of neuronal transcripts. *Nucleic Acids Res.* **2018**, *46*, 6561–6575. [[CrossRef](#)]
80. Lee, J.; Kim, M.; Itoh, T.Q.; Lim, C. Ataxin-2: A versatile posttranscriptional regulator and its implication in neural function. *Wiley Interdiscip. Rev. RNA* **2018**, *9*, 1–13. [[CrossRef](#)]
81. Lastres-Becker, I.; Rüb, U.; Auburger, G. Spinocerebellar ataxia 2 (SCA2). *Cerebellum* **2008**, *7*, 115–124. [[CrossRef](#)] [[PubMed](#)]
82. Becker, L.A.; Huang, B.; Bieri, G.; Ma, R.; Knowles, D.A.; Jafar-Nejad, P.; Messing, J.; Kim, H.J.; Soriano, A.; Auburger, G.; et al. Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice. *Nature* **2017**, *544*, 367–371. [[CrossRef](#)]

83. Tian, B.; Hu, J.; Zhang, H.; Lutz, C.S. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res.* **2005**, *33*, 201–212. [[CrossRef](#)]
84. Lianoglou, S.; Garg, V.; Yang, J.L.; Leslie, C.S.; Mayr, C. Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes Dev.* **2013**, *27*, 2380–2396. [[CrossRef](#)] [[PubMed](#)]
85. Macdonald, C.C. Tissue-specific mechanisms of alternative polyadenylation: Testis, brain, and beyond (2018 update). *RNA* **2019**, *10*, e1526. [[CrossRef](#)]
86. Smibert, P.; Miura, P.; Westholm, J.O.; Shenker, S.; May, G.; Duff, M.O.; Zhang, D.; Eads, B.D.; Carlson, J.; Brown, J.B.; et al. Global Patterns of Tissue-Specific Alternative Polyadenylation in *Drosophila*. *Cell Rep.* **2012**, *1*, 277–289. [[CrossRef](#)]
87. Miura, P.; Shenker, S.; Andreu-Agullo, C.; Westholm, J.O.; Lai, E.C. Widespread and extensive lengthening of 3' UTRs in the mammalian brain. *Genome Res.* **2013**, *23*, 812–825. [[CrossRef](#)]
88. Wei, L.; Lee, S.; Majumdar, S.; Zhang, B.; Sanfilippo, P.; Joseph, B.; Miura, P.; Soller, M.; Lai, E.C. Overlapping Activities of ELAV/Hu Family RNA Binding Proteins Specify the Extended Neuronal 3' UTR Landscape in *Drosophila*. *Mol. Cell* **2020**, *80*, 140–155.e6. [[CrossRef](#)]
89. Soller, M.; White, K. ELAV inhibits 3'-end processing to promote neural splicing of ewg pre-mRNA. *Genes Dev.* **2003**, *17*, 2526–2538. [[CrossRef](#)] [[PubMed](#)]
90. Bae, B.; Miura, P. Emerging roles for 3' UTRs in neurons. *Int. J. Mol. Sci.* **2020**, *21*, 3413. [[CrossRef](#)]
91. Hinman, M.N.; Lou, H. Diverse molecular functions of Hu proteins. *Cell. Mol. Life Sci.* **2008**, *65*, 3168–3181. [[CrossRef](#)]
92. Fernandopulle, M.S.; Lippincott-Schwartz, J.; Ward, M.E. RNA transport and local translation in neurodevelopmental and neurodegenerative disease. *Nat. Neurosci.* **2021**, *24*, 622–632. [[CrossRef](#)] [[PubMed](#)]
93. Blazie, S.M.; Babb, C.; Wilky, H.; Rawls, A.; Park, J.G.; Mangone, M. Comparative RNA-Seq analysis reveals pervasive tissue-specific alternative polyadenylation in *Caenorhabditis elegans* intestine and muscles. *BMC Biol.* **2015**, *13*, 4. [[CrossRef](#)] [[PubMed](#)]
94. Relaix, F.; Rocancourt, D.; Mansouri, A.; Buckingham, M. Divergent functions of murine Pax3 and Pax7 in limb muscle development. *Genes Dev.* **2004**, *18*, 1088–1105. [[CrossRef](#)]
95. Der Vartanian, A.; Quéting, M.; Michineau, S.; Auradé, F.; Hayashi, S.; Dubois, C.; Rocancourt, D.; Drayton-Libotte, B.; Szegedi, A.; Buckingham, M.; et al. PAX3 Confers Functional Heterogeneity in Skeletal Muscle Stem Cell Responses to Environmental Stress. *Cell Stem Cell* **2019**, *24*, 958–973.e9. [[CrossRef](#)] [[PubMed](#)]
96. Scaramozza, A.; Park, D.; Kollu, S.; Beerman, I.; Sun, X.; Rossi, D.J.; Lin, C.P.; Scadden, D.T.; Crist, C.; Brack, A.S. Lineage Tracing Reveals a Subset of Reserve Muscle Stem Cells Capable of Clonal Expansion under Stress. *Cell Stem Cell* **2019**, *24*, 944–957.e5. [[CrossRef](#)] [[PubMed](#)]
97. Boutet, S.C.; Cheung, T.H.; Quach, N.L.; Liu, L.; Prescott, S.L.; Edalati, A.; Iori, K.; Rando, T.A. Alternative polyadenylation mediates microRNA regulation of muscle stem cell function. *Cell Stem Cell* **2012**, *10*, 327–336. [[CrossRef](#)] [[PubMed](#)]
98. Hirai, H.; Verma, M.; Watanabe, S.; Tastad, C.; Asakura, Y.; Asakura, A. MyoD regulates apoptosis of myoblasts through microRNA-mediated down-regulation of Pax3. *J. Cell Biol.* **2010**, *191*, 347–365. [[CrossRef](#)]
99. Crist, C.G.; Montarras, D.; Pallafacchina, G.; Rocancourt, D.; Cumano, A.; Conway, S.J.; Buckingham, M. Muscle stem cell behavior is modified by microRNA-27 regulation of Pax3 expression. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13383–13387. [[CrossRef](#)]
100. Tomé, F.M.S.; Chateau, D.; Helbling-Leclerc, A.; Fardeau, M. Morphological changes in muscle fibers in oculopharyngeal muscular dystrophy. *Neuromuscul. Disord.* **1997**, *7*, 63–69. [[CrossRef](#)]
101. Malerba, A.; Klein, P.; Bachtarzi, H.; Jarmin, S.A.; Cordova, G.; Ferry, A.; Strings, V.; Espinoza, M.P.; Mamchaoui, K.; Blumen, S.C.; et al. PABPN1 gene therapy for oculopharyngeal muscular dystrophy. *Nat. Commun.* **2017**, *8*, 14848. [[CrossRef](#)]
102. Apponi, L.H.; Leung, S.W.; Williams, K.R.; Valentini, S.R.; Corbett, A.H.; Pavlath, G.K. Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis. *Hum. Mol. Genet.* **2010**, *19*, 1058–1065. [[CrossRef](#)] [[PubMed](#)]
103. Simonelig, M. PABPN1 shuts down alternative poly(A) sites. *Cell Res.* **2012**, *22*, 1419–1421. [[CrossRef](#)]
104. Jacquemont, S.; Hagerman, R.J.; Hagerman, P.J.; Leehey, M.A. Fragile-X syndrome and fragile X-associated tremor/ataxia syndrome: Two faces of FMR1. *Lancet Neurol.* **2007**, *6*, 45–55. [[CrossRef](#)]
105. Sunamura, N.; Iwashita, S.; Enomoto, K.; Kadoshima, T.; Isono, F. Loss of the fragile X mental retardation protein causes aberrant differentiation in human neural progenitor cells. *Sci. Rep.* **2018**, *8*, 11585. [[CrossRef](#)] [[PubMed](#)]
106. Cabal-Herrera, A.M.; Tassanakijpanich, N.; Salcedo-Arellano, M.J.; Hagerman, R.J. Fragile X-Associated Tremor/Ataxia Syndrome (FXTAS): Pathophysiology and Clinical Implications. *Int. J. Mol. Sci.* **2020**, *21*, 4391. [[CrossRef](#)] [[PubMed](#)]
107. Tassone, F.; De Rubeis, S.; Carosi, C.; La Fata, G.; Serpa, G.; Raske, C.; Willemsen, R.; Hagerman, P.J.; Bagni, C. Differential usage of transcriptional start sites and polyadenylation sites in FMR1 premutation alleles. *Nucleic Acids Res.* **2011**, *39*, 6172–6185. [[CrossRef](#)]
108. Melamed, Z.; Lopez-Erauskin, J.; Baughn, M.W.; Zhang, O.; Drenner, K.; Sun, Y.; Freyermuth, F.; McMahan, M.A.; Beccari, M.S.; Artates, J.; et al. Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. *Nat. Neurosci.* **2019**, *22*, 180. [[CrossRef](#)]
109. Patel, R.; Brophy, C.; Hickling, M.; Neve, J.; Furger, A. Alternative cleavage and polyadenylation of genes associated with protein turnover and mitochondrial function are deregulated in Parkinson's, Alzheimer's and ALS disease. *BMC Med. Genom.* **2019**, *12*, 1–14. [[CrossRef](#)]
110. Balendra, R.; Isaacs, A.M. C9orf72-mediated ALS and FTD: Multiple pathways to disease. *Nat. Rev. Neurol.* **2018**, *14*, 544–558. [[CrossRef](#)]

111. Niblock, M.; Smith, B.N.; Lee, Y.B.; Sardone, V.; Topp, S.; Troakes, C.; Al-Sarraj, S.; Leblond, C.S.; Dion, P.A.; Rouleau, G.A.; et al. Retention of hexanucleotide repeat-containing intron in C9orf72 mRNA: Implications for the pathogenesis of ALS/FTD. *Acta Neuropathol. Commun.* **2016**, *4*, 18. [[CrossRef](#)]
112. Prudencio, M.; Belzil, V.V.; Batra, R.; Ross, C.A.; Gendron, T.F.; Pregent, L.J.; Murray, M.E.; Overstreet, K.K.; Piazza-Johnston, A.E.; Desaro, P.; et al. Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat. Neurosci.* **2015**, *18*, 1175–1182. [[CrossRef](#)] [[PubMed](#)]
113. Giannini, M.; Bayona-Feliu, A.; Sproviero, D.; Barroso, S.I.; Cereda, C.; Aguilera, A. TDP-43 mutations link Amyotrophic Lateral Sclerosis with R-loop homeostasis and R loop-mediated DNA damage. *PLoS Genet.* **2020**, *16*, e1009260. [[CrossRef](#)] [[PubMed](#)]
114. Prashad, S.; Gopal, P.P. RNA-binding proteins in neurological development and disease. *RNA Biol.* **2021**, *18*, 972. [[CrossRef](#)] [[PubMed](#)]
115. Humphrey, J.; Birsa, N.; Milioto, C.; McLaughlin, M.; Ule, A.M.; Robaldo, D.; Eberle, A.B.; Kräuchi, R.; Bentham, M.; Brown, A.L.; et al. FUS ALS-causative mutations impair FUS autoregulation and splicing factor networks through intron retention. *Nucleic Acids Res.* **2020**, *48*, 6889–6905. [[CrossRef](#)] [[PubMed](#)]
116. Koyama, A.; Sugai, A.; Kato, T.; Ishihara, T.; Shiga, A.; Toyoshima, Y.; Koyama, M.; Konno, T.; Hirokawa, S.; Yokoseki, A.; et al. Increased cytoplasmic TARDBP mRNA in affected spinal motor neurons in ALS caused by abnormal autoregulation of TDP-43. *Nucleic Acids Res.* **2016**, *44*, 5820–5836. [[CrossRef](#)] [[PubMed](#)]
117. LoRusso, S.; Weiner, B.; Arnold, W.D. Myotonic Dystrophies: Targeting Therapies for Multisystem Disease. *Neurotherapeutics* **2018**, *15*, 872–884. [[CrossRef](#)]
118. Ozimski, L.L.; Sabater-Arcis, M.; Bargiela, A.; Artero, R. The hallmarks of myotonic dystrophy type 1 muscle dysfunction. *Biol. Rev.* **2021**, *96*, 716–730. [[CrossRef](#)]
119. Kurkiewicz, A.; Cooper, A.; McIlwaine, E.; Cumming, S.A.; Adam, B.; Krahe, R.; Puymirat, J.; Schoser, B.; Timchenko, L.; Ashizawa, T.; et al. Towards development of a statistical framework to evaluate myotonic dystrophy type 1 mRNA biomarkers in the context of a clinical trial. *PLoS ONE* **2020**, *15*, e0231000. [[CrossRef](#)]
120. Gudde, A.E.E.G.; van Kessel, I.D.G.; André, L.M.; Wieringa, B.; Wansink, D.G. Trinucleotide-repeat expanded and normal DMPK transcripts contain unusually long poly(A) tails despite differential nuclear residence. *Biochim. Biophys. Acta—Gene Regul. Mech.* **2017**, *1860*, 740–749. [[CrossRef](#)]
121. Kino, Y.; Mori, D.; Oma, Y.; Takeshita, Y.; Sasagawa, N.; Ishiura, S. Muscleblind protein, MBNL1/EXP, binds specifically to CHHG repeats. *Hum. Mol. Genet.* **2004**, *13*, 495–507. [[CrossRef](#)]
122. Yuan, Y.; Compton, S.A.; Sobczak, K.; Stenberg, M.G.; Thornton, C.A.; Griffith, J.D.; Swanson, M.S. Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. *Nucleic Acids Res.* **2007**, *35*, 5474–5486. [[CrossRef](#)] [[PubMed](#)]
123. Konieczny, P.; Stepniak-Konieczna, E.; Sobczak, K. MBNL proteins and their target RNAs, interaction and splicing regulation. *Nucleic Acids Res.* **2014**, *42*, 10873–10887. [[CrossRef](#)] [[PubMed](#)]
124. Taylor, K.; Sznajder, Ł.J.; Cywoniuk, P.; Thomas, J.D.; Swanson, M.S.; Sobczak, K. MBNL splicing activity depends on RNA binding site structural context. *Nucleic Acids Res.* **2018**, *46*, 9119–9133. [[CrossRef](#)]
125. Brinegar, A.E.; Cooper, T.A. Roles for RNA-binding proteins in development and disease. *Brain Res.* **2016**, *1647*, 1–8. [[CrossRef](#)] [[PubMed](#)]
126. Wang, E.T.; Cody, N.A.L.; Jog, S.; Biancolella, M.; Wang, T.T.; Treacy, D.J.; Luo, S.; Schroth, G.P.; Housman, D.E.; Reddy, S.; et al. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* **2012**, *150*, 710–724. [[CrossRef](#)] [[PubMed](#)]
127. Itskovich, S.S.; Gurunathan, A.; Clark, J.; Burwinkel, M.; Wunderlich, M.; Berger, M.R.; Kulkarni, A.; Chetal, K.; Venkatasubramanian, M.; Salomonis, N.; et al. MBNL1 regulates essential alternative RNA splicing patterns in MLL-rearranged leukemia. *Nat. Commun.* **2020**, *11*, 2369. [[CrossRef](#)]
128. Mankodi, A. Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum. Mol. Genet.* **2001**, *10*, 2165–2170. [[CrossRef](#)]
129. Jiang, H.; Mankodi, A.; Swanson, M.S.; Moxley, R.T.; Thornton, C.A. Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum. Mol. Genet.* **2004**, *13*, 3079–3088. [[CrossRef](#)]
130. Song, K.Y.; Guo, X.M.; Wang, H.Q.; Zhang, L.; Huang, S.Y.; Huo, Y.C.; Zhang, G.; Feng, J.Z.; Zhang, R.R.; Ma, Y.; et al. MBNL1 reverses the proliferation defect of skeletal muscle satellite cells in myotonic dystrophy type 1 by inhibiting autophagy via the mTOR pathway. *Cell Death Dis.* **2020**, *11*, 545. [[CrossRef](#)]
131. MacDonald, M.E.; Ambrose, C.M.; Duyao, M.P.; Myers, R.H.; Lin, C.; Srinidhi, L.; Barnes, G.; Taylor, S.A.; James, M.; Groot, N.; et al. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **1993**, *72*, 971–983. [[CrossRef](#)]
132. Tabrizi, S.J.; Flower, M.D.; Ross, C.A.; Wild, E.J. Huntington disease: New insights into molecular pathogenesis and therapeutic opportunities. *Nat. Rev. Neurol.* **2020**, *16*, 529–546. [[CrossRef](#)]
133. Mccolgan, P.; Tabrizi, S.J. Huntington's disease: A clinical review. *Eur. J. Neurol.* **2018**, *25*, 24–34. [[CrossRef](#)] [[PubMed](#)]
134. Romo, L.; Ashar-Patel, A.; Pfister, E.; Aronin, N. Alterations in mRNA 3' UTR Isoform Abundance Accompany Gene Expression Changes in Human Huntington's Disease Brains. *Cell Rep.* **2017**, *20*, 3057–3070. [[CrossRef](#)]

135. Lin, B.; Rommens, J.M.; Graham, R.K.; Kalchman, M.; Macdonald, H.; Nasir, J.; Delaney, A.; Goldberg, Y.P.; Hayden, M.R. Differential 3' polyadenylation of the huntington disease gene results in two mRNA species with variable tissue expression. *Hum. Mol. Genet.* **1993**, *2*, 1541–1545. [[CrossRef](#)] [[PubMed](#)]
136. Xu, H.; An, J.J.; Xu, B. Distinct cellular toxicity of two mutant huntingtin mRNA variants due to translation regulation. *PLoS ONE* **2017**, *12*, e0177610. [[CrossRef](#)] [[PubMed](#)]
137. Sathasivam, K.; Neueder, A.; Gipson, T.A.; Landles, C.; Benjamin, A.C.; Bondulich, M.K.; Smith, D.L.; Faull, R.L.M.; Roos, R.A.C.; Howland, D.; et al. Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 2366–2370. [[CrossRef](#)] [[PubMed](#)]
138. Neueder, A.; Landles, C.; Ghosh, R.; Howland, D.; Myers, R.H.; Faull, R.L.M.; Tabrizi, S.J.; Bates, G.P. The pathogenic exon 1 HTT protein is produced by incomplete splicing in Huntington's disease patients. *Sci. Rep.* **2017**, *7*, 1307. [[CrossRef](#)]
139. Neueder, A.; Dumas, A.A.; Benjamin, A.C.; Bates, G.P. Regulatory mechanisms of incomplete huntingtin mRNA splicing. *Nat. Commun.* **2018**, *9*, 3955. [[CrossRef](#)]
140. Mason, M.A.; Gomez-Paredes, C.; Sathasivam, K.; Neueder, A.; Papadopoulou, A.S.; Bates, G.P. Silencing Srsf6 does not modulate incomplete splicing of the huntingtin gene in Huntington's disease models. *Sci. Rep.* **2020**, *10*, 14057. [[CrossRef](#)]
141. Lieberman, A.P.; Shakkottai, V.G.; Albin, R.L. Polyglutamine Repeats in Neurodegenerative Diseases. *Annu. Rev. Pathol. Mech. Dis.* **2019**, *14*, 1–27. [[CrossRef](#)]
142. Kamieniarz-Gdula, K.; Proudfoot, N.J. Transcriptional Control by Premature Termination: A Forgotten Mechanism. *Trends Genet.* **2019**, *35*, 553–564. [[CrossRef](#)]
143. Ciesiolka, A.; Jazurek, M.; Drazkowska, K.; Krzyzosiak, W.J. Structural characteristics of simple RNA repeats associated with disease and their deleterious protein interactions. *Front. Cell. Neurosci.* **2017**, *11*, 1–19. [[CrossRef](#)]
144. Howe, K.L.; Achuthan, P.; Allen, J.; Allen, J.; Alvarez-Jarreta, J.; Ridwan Amode, M.; Armean, I.M.; Azov, A.G.; Bennett, R.; Bhai, J.; et al. Ensembl 2021. *Nucleic Acids Res.* **2021**, *49*, D884–D891. [[CrossRef](#)]
145. Chen, C.Y.; Chen, S.T.; Juan, H.F.; Huang, H.C. Lengthening of 3'UTR increases with morphological complexity in animal evolution. *Bioinformatics* **2012**, *28*, 3178–3181. [[CrossRef](#)]
146. Mayr, C. What are 3' utrs doing? *Cold Spring Harb. Perspect. Biol.* **2019**, *11*. [[CrossRef](#)]
147. Zhu, S.; Lian, Q.; Ye, W.; Qin, W.; Wu, Z.; Ji, G.; Wu, X. scAPAdb: A comprehensive database of alternative polyadenylation at single-cell resolution. *Nucleic Acids Res.* **2021**, gkab795. [[CrossRef](#)]
148. Gao, Y.; Li, L.; Amos, C.I.; Li, W. Analysis of alternative polyadenylation from singlecell RNA-seq using scDaPars reveals cell subpopulations invisible to gene expression. *Genome Res.* **2021**, *31*, 1856–1866. [[CrossRef](#)]
149. Saxena, S.; Caroni, P. Selective Neuronal Vulnerability in Neurodegenerative Diseases: From Stressor Thresholds to Degeneration. *Neuron* **2011**, *71*, 35–48. [[CrossRef](#)]
150. Ogorodnikov, A.; Levin, M.; Tattikota, S.; Tokalov, S.; Hoque, M.; Scherzinger, D.; Marini, F.; Poetsch, A.; Binder, H.; Macher-Göppinger, S.; et al. Transcriptome 3' end organization by PCF11 links alternative polyadenylation to formation and neuronal differentiation of neuroblastoma. *Nat. Commun.* **2018**, *9*, 5331. [[CrossRef](#)]
151. Masamha, C.P.; Xia, Z.; Yang, J.; Albrecht, T.R.; Li, M.; Shyu, A.-B.; Li, W.; Wagner, E.J. CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. *Nature* **2014**, *510*, 412–416. [[CrossRef](#)]
152. Fiszer, A.; Krzyzosiak, W.J. Oligonucleotide-based strategies to combat polyglutamine diseases. *Nucleic Acids Res.* **2014**, *42*, 6787–6810. [[CrossRef](#)] [[PubMed](#)]
153. Fiszer, A.; Olejniczak, M.; Galka-Marciniak, P.; Mykowska, A.; Krzyzosiak, W.J. Self-duplexing CUG repeats selectively inhibit mutant huntingtin expression. *Nucleic Acids Res.* **2013**, *41*, 10426–10437. [[CrossRef](#)]
154. Ciesiolka, A.; Stroynowska-Czerwinska, A.; Joachimiak, P.; Ciolak, A.; Kozłowska, E.; Michalak, M.; Dabrowska, M.; Olejniczak, M.; Raczynska, K.D.; Zielinska, D.; et al. Artificial miRNAs targeting CAG repeat expansion in ORFs cause rapid deadenylation and translation inhibition of mutant transcripts. *Cell. Mol. Life Sci.* **2020**. [[CrossRef](#)]